Form PTO-1390 (REV. 11-2000) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

ATTORNEY'S DOCKET NUMBER 9013-44

U.S. APPLICATION NO. (If known, see 37 CFR 1.5) 10/070496

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/GB00/03444	September 6, 2000	September 7, 1999

TITLE OF INVENTION
ESSENTIAL GENES AND ASSAYS RELATING THERETO

APPLICANT(S) FOR DO/EO/US
Roger Wayne Davies, Kim Kaiser, Ming Yao Yang

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☐ This is an express request to begin national examination procedures (35 U.S.C. 371(f)). The submission must include items (5), (6), (9) and (21) indicated below.
4. ☒ The US has been elected by the expiration of 19 months from the priority date (Article 31).
5. ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. ☒ is attached hereto (required only if not communicated by the International Bureau).
   b. ☐ has been communicated by the International Bureau.
   c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).
   a. ☐ is attached hereto.
   b. ☐ has been previously submitted under 35 U.S.C. 154(d)(4)
7. ☒ Amendments to the claims of the International Application Under PCT Article 19 (35 U.S.C. 371(c)(3))
   a. ☐ are attached hereto (required only if not communicated by the International Bureau).
   b. ☐ have been communicated by the International Bureau.
   c. ☒ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☐ have not been made and will not be made.
8. ☐ An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☐ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. ☐ An English language translation of the annexes of the International Preliminary Examination Report Under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11 to 20 below concern document(s) or information included:

11. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. ☒ A FIRST preliminary amendment.
14. ☐ A SECOND or SUBSEQUENT preliminary amendment.
15. ☐ A substitute specification.
16. ☐ A change of power of attorney and/or address letter.
17. ☐ A computer-readable form of the sequence listing in accordance with PCT Rule 13ter.2 and 35 U.S.C. 1.821 - 1.825.
18. ☐ A second copy of the published international application under 35 U.S.C. 154(d)(4).
19. ☐ A second copy of the English language translation of the international application under 35 U.S.C. 154(d)(4)
20. ☒ Other items or information: International Preliminary Examination Report, PCT Request

U.S. APPLICATION NO (if known, see 37 CFR 1.5) 10/070498	INTERNATIONAL APPLICATION NO PCT/GB00/03444	ATTORNEY DOCKET NO 9013-44

21. ☐ The following fees are submitted:

BASIC NATIONAL FEE (37 CFR 1.492(a) (1) - (5)):	CALCULATIONS	PTO USE ONLY
Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO **$1040.00**		
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO..... **$890.00**	890.00	
International preliminary examination fee 37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO..... **$740.00**		
International preliminary examination fee (37 CFR 1.482) paid to USPTO but all claims did not satisfy provisions of PCT Article 33(1)-(4).. ..... **$710.00**		
International preliminary examination fee (37 CFR 1.482) paid to USPTO and all claims satisfied provisions of PCT Article 33(1)-(4)..... **$100.00**		
ENTER APPROPRIATE BASE FEE AMOUNT =	$890.00	
Surcharge of **$130.00** for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(e)).	$	

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE	
Total claims	31 - 20 =	11	x **$18.00**	$198.00
Independent Claims	11 - 3 =	8	x **$84.00**	$672.00
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ **$280.00**	$
TOTAL OF ABOVE CALCULATIONS =				$1760.00
☐ Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2			-	$
SUBTOTAL =				$
Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(f)).			+	$
TOTAL NATIONAL FEE =				$
Fee for Recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). **$40.00** per property			+	$
TOTAL FEES ENCLOSED =				$1760.00
			Amount to be refunded:	$
			charged:	$

a. ☒ A check in the amount of $1760.00 to cover the above fees is enclosed.

b. ☐ Please charge my Deposit Account No. 50-0220 in the amount of $_____ to cover the above fees. A duplicate copy of this sheet is enclosed.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 50-0220. A duplicate copy of this sheet is enclosed.

d. ☐ Fees are to be charged to a credit card. **WARNING:** Information on this form may become public. **Credit card information should not be included on this form.** Provide credit card information and authorization on PTO-2038.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

20792
PATENT TRADEMARK OFFICE

P. Michael Sajovec, Reg. No. 31,793

Date: March 7, 2002

CERTIFICATE OF EXPRESS MAILING

Express Mail Label No. EV015665007US
Date of Deposit: March 7, 2002

I hereby certify that this correspondence is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to: BOX PCT, Attn: DO/EO/US, Commissioner for Patents, Washington, DC 20231.

Monica L. Croom

Attorney's Docket No. 9013-44 PATENT

IN THE UNITED STATES DESIGNATED OFFICE (DO/US)

In re: Davies et al.
Serial No. to be assigned
Filed: concurrently herewith
For: *ESSENTIAL GENES AND ASSAYS RELATING THERETO*

Date: March 7, 2002

BOX PCT
Commissioner for Patents
Washington, DC 20231

PRELIMINARY AMENDMENT

Sir:

Prior to the examination of the above application, please amend the above-identified application as follows.

IN THE SPECIFICATION:

Please amend the specification as follows:

At page one, following the title "ESSENTIAL GENES AND ASSAYS RELATING THERETO", please insert:

-- **Related Applications**

This application claims priority under 35 U.S.C. § 371 from PCT Application No. PCT/GB00/03444, filed in English on September 6, 2000, which claims the benefit of Great Britain Application Serial No. 9921009.8, filed on September 7, 1999, the disclosures of which are incorporated by reference herein in their entireties.--

IN THE CLAIMS:

Please cancel Claims 1-36.

Please enter the following new claims:

37. A method of identifying genes suitable for use in the development of pesticides or cancer therapies, the method comprising the steps:

providing Drosophila fly lines which display a lethal or semi-lethal phenotype having been generated using the technique of P-element transposon-tagged insertion;

conducting plasmid rescue so as to isolate nucleic acid surrounding the site of transposon insertion; and

utilizing the isolated nucleic acid to clone a larger portion of nucleic acid containing a complete essential gene.

38. A method of identifying genes suitable for use in the development of pesticide or cancer therapies, the method comprising the steps of:

providing Drosophila fly lines which display a lethal or semi-lethal phenotype having been generated using the technique of P-element transposon-tagged insertion;

conducting plasmid rescue so as to isolate nucleic acid surrounding the site of transposon insertion;

utilising the isolated nucleic acid to clone a larger portion of nucleic acid containing a complete essential gene; and

utilising the larger portion of nucleic acid in hybridisation studies to identify an essential gene from another organism such as a mammal.

39. A screening assay for identifying compounds which have a physiological effect on an organism, the assay comprising the steps of:

a) reacting a test compound with a protein encoded by an essential gene comprising a sequence selected from the group consisting of SEQ ID Nos. 1-902, specific fragment thereof, or homologue thereof, from the organism; and

b) detecting any modulatory effect the compound has on the protein.

40. A screening assay for identifying compounds which have a physiological

effect on an organism, the assay comprising the steps of:

a) reacting a test compound with a protein encoded by an essential gene comprising a sequence selected from the group consisting of SEQ ID Nos. 430-783 and 899-902, specific fragment thereof, or homologue thereof, from the organism; and

b) detecting any modulatory effect the compound has on the protein.

41. The screening assay according to claim 39 wherein the effect on the protein is a negative modulation.

42. The screening assay according to claim 39 wherein the assay is a ligand binding assay for detecting the effect the compound has on the ligand binding of the protein

43. The screening assay according to claim 39 wherein the assay is a functional activity assay for detecting the effect the compound has on the functional activity of the protein.

44. The screening assay according to claim 43 wherein the functional activity assay is selected from the group consisting of kinase assays; protein phosphatase assays; adenyl cyclase assays; guanylyl cyclase assays; phosphodiesterase assays; nucleosidease assays; protease assays; protein secretion and/or import assays; nuclease assays; DNA metabolism assays; transcription factor assays; apoptosis assays; calcium utilisation assays; receptor/ion channel assays; and G protein assays.

45. A compound having modulatory activity on a protein encoded by an essential gene, as identified by an assay according to claim 39.

46. A pesticidal formulation comprising a compound according to claim 45, together with a pesticidally acceptable excipient.

47. A pesticidally active compound identified by an assay according to claim 37 and further tested for its ability to kill pests.

48. A method of selectively modulating activity, in an organism, of a protein encoded by an essential gene comprising a sequence selected from the group consisting of SEQ ID Nos. 1-902 or a specific fragment thereof, or homologue thereof, comprising administering a compound that selectively modulates activity of the protein in the organism.

49. The method according to claim 48, wherein the selective modulation in activity of the protein has the result of substantially eliminating or severely reducing the activity of the protein, as compared to the activity of the protein without modulation.

50. The method according to claims 48, wherein the compound modulates the activity of the protein and has a minimal modulatory effect on other proteins of the organism.

51. The method according to claim 46, wherein the modulation in activity of the protein has the effect of being lethal or semi-lethal to the organism.

52. A method of modulating activity, in an organism, of a protein encoded by an essential gene comprising a sequence selected from the group consisting of SEQ ID Nos. 1-902 or a specific fragment thereof, or homologue thereof, comprising administering a compound, that selectively modulates activity of the protein, to an organism, and wherein the ability of the protein to modulate the activity of the protein is determined by:
- exposing the protein which has been produced by a genetically engineered cell expressing the protein, with the compound for a period of time;
- measuring the activity of the exposed protein using a ligand binding or functional activity assay; and

- comparing the activity of the exposed protein with an activity of a control protein which has not been exposed to the compound, so that compounds that modulate the protein activity are identified.

53. The method according to claim 52, for selectively modulating activity, in an organism, of a protein, further comprising the steps of:
- exposing a further cellular protein(s) of the organism to the compound for a period of time;
- measuring the activity of said further protein(s) using an assay(s) appropriate for such a purpose; and
- comparing the activity of said exposed further cellular protein(s) with an activity of a control protein(s) which has not been exposed to the compound, so that compounds that substantially do not, or minimally modulate said further cellular protein(s) activity, are identified.

54. A method of identifying compounds having a potentially pesticidal activity caused by modulation of a protein encoded by an essential gene comprising a sequence selected from the group consisting of SEQ ID Nos. 1-902 or a specific fragment thereof, or homologue thereof, which comprises;

- obtaining the protein by heterologous expression of the essential gene in a host cell;
- employing the protein in an assay according to claim 39 for detecting a compound which displays modulatory activity on the protein; and
- testing the compound which displays modulatory activity on the protein for its pesticidal activity on an organism.

55. A compound identified by the method according to claim 54 as having pesticidal activity.

56. A pesticidal formulation comprising a compound according to claim 55

identified as having pesticidal activity, together with a pesticidally acceptable excipient

57. A method for the production of a pesticidal composition comprising identifying a compound that displays pesticidal activity using the method according to claim 54 and mixing the compound identified, or a derivative, or an analogue thereof, with a pesticidally acceptable carrier.

58. An isolated polynucleotide fragment comprising a sequence selected from the group consisting of SEQ ID Nos.430-783 and 899-902, a fragment thereof, or a homologue thereof.

59. An essential gene comprising a sequence selected from the group consisting of SEQ ID Nos.430-783 and 899-902, a fragment thereof, or a homologue thereof.

60. An isolated polynucleotide which hybridises under stringent conditions to a polynucleotide fragment selected from the group consisting of SEQ ID Nos. 430-783 and 899-902 or a fragment thereof.

61. An essential gene comprising a sequence selected from the group consisting of SEQ ID Nos.430-783 and 899-902, a fragment thereof, or a homologue thereof.

62. An expression vector comprising the essential gene according to claim 61.

63. An expression vector according to claim 62 comprising one or more control sequences capable of directing the replication and/or expression of an operatively linked essential gene.

64. A prokaryotic or eukaryotic host cell comprising the expression vector

according to claim 62.

65. A method of producing a polypeptide comprising culturing a host cell according to claim 64 under conditions permitting expression of the polypeptide.

66. A polypeptide produced by the method of claim 65.

67. A method of identifying and facilitating isolation of an essential gene from an organism, comprising the steps of:

providing a polynucleotide fragment comprising a sequence selected from the group consisting of SEQ ID Nos. 430-783 and 899-902, or a fragment thereof; and

allowing the polynucleotide fragment to specifically hybridise to nucleic acid from the organism such that a corresponding essential gene from the organism is identified.

REMARKS

The above amendment to the specification has been made to claim priority to the identified PCT and British patent applications. The above claims have been amended to better conform to U.S. practice. Applicants respectfully request substantive examination on the merits.

Respectfully submitted,

F. Michael Sajovec
Registration No. 31,793

In re: Davies et al.
Serial No. to be assigned
Filed: concurrently herewith
Page 8 of 8

CERTIFICATE OF EXPRESS MAILING

Express Mail Label No. EV015665007US
Date of Deposit: March 7, 2002

I hereby certify that this correspondence is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to: BOX PCT, Attn: DO/EO/US, Commissioner for Patents, Washington, DC 20231.

Monica L. Croom

Monica L. Croom

ESSENTIAL GENES AND ASSAYS RELATING THERETO

The present invention relates in part to target based screening assays, particularly for pesticides, based on the identification and use of essential genes/proteins, as well as novel genes/proteins themselves and compounds identified by such assays which may have a modulatory effect on the protein. It also relates to screening assays for compounds with therapeutic use in cancer therapy or other proliferative diseases.

Crop destruction by organisms (eg. pests) such as insects results in a considerable economic loss and serious reduction in productivity. Chemical pesticides are typically used in order to control the pests and reduce crop loss. However pesticide development has generally been less than controlled or focussed, such that the biochemical or genetical functions of the pesticide have not been a major concern, but rather simply whether or not the pesticide was effective ie. killed the pests.

However, increasing environmental concerns and development of resistance to existing pesticides has led to a more rational approach to pesticide development being voiced.

It is amongst the objects of the present invention to provide a more rational approach to pesticide development by providing pesticide screening assays based on the identification and use of genes/proteins which are considered to be essential to the pest.

The present inventors have used genetic techniques in order to study a model "pest", namely *Drosophila*. As a result of these studies, the present inventors have identified a considerable number of essential genes/proteins, which may be used in assays based on the functional activity of the protein and/or ligand binding assays for screening for modulators of protein activity which have potential use for example as pesticides.

Thus, in a first aspect the present invention provides a screening assay for identifying compounds which have a physiological or biochemical effect on an organism, the assay comprising the steps of:

a) reacting a test compound with a protein encoded by an essential gene comprising a sequence selected from the group consisting of SEQ ID Nos. 1-902, specific fragment thereof, or homologue thereof, from the organism; and

b) detecting any modulatory effect the compound has on the protein.

A modulatory effect is one which alters the function and/or activity of the protein. For example, negative modulation, as used herein, refers to a reduction in the level and/or activity of an essential gene product (eg. polypeptide) relative to the level and/or activity of the essential gene product in absence of the modulatory treatment. The reduction in the level and/or activity may be taken to be for example less than 50%, 40%, 30%, 20%, 10%, 5% or 1% of the level and/or activity of the essential gene product in the absence of the modulatory treatment. Positive modulation, as used herein, refers to an increase in the level and/or activity of the essential gene product, relative to the level and/or activity of the gene product in the absence of the modulatory treatment. For example greater than 150%, 200%, 300%, 400% increase with respect to the relative level in the absence of the modulatory treatment.

Typically the assay is designed to be used to screen for compounds which effect the physiology or biochemistry in a manner which is harmful or biocidal eg. pesticidal, to the organism. However, the assay may be used to screen for other effects, such as beneficial or therapeutic effects. For example the assay may be used to screen for compounds which are effective against cancer or other proliferative diseases. In this manner it may be possible to identify proteins which are negatively or positively modulated in a patient suffering from for example cancer. Consequently

the assay would be used to identify compounds which may reduce or substantially eliminate such negative or positive modulation.

An essential gene is one for which it has been determined that the expression of a functional protein is necessary in order to avoid a lethal or semi-lethal phenotype. A lethal phenotype is defined as a phenotype characterised by organism death due to cellular or system failure at some developmental stage. A semi-lethal phenotype is for example characterised by low fecundity (the tendency to produce none or only few offspring, eg. less than 50%, 40%, 30%, 20%, 10%, 5% or 1% compared to wildtype), and frequently by short lifespan, or a behavioural modification which leads to reduced ability to generate offspring.

Studies carried out by the present inventors have identified many distinct fly lines, in the model organism, *Drosophila*, which display a lethal or semi-lethal phenotype. Fly lines which display such a lethal or semi-lethal phenotype have been generated using the technique of P-element transposon-tagged insertion (1) Torok, T,G. Tick, M. Alvarado and I. Kiss, 1993 P-lacW Insertional Mutagenesis on the second chromosome of Drosophila melanogaster: Isolation of lethals with different overgrowth phenotypes. Genetics 135: 71 - 80; (2) Dèak, P., M.M. Omar, R.D.C. Saunders, M. Pal, O. Komonyi, J. Szidonya, P. Maroy, Y. Zhang, M. Ashburner, P. Benos, C. Savakis, I. Siden-Kiamos, C. Louis, V.N. Bolshakov, F.C. Kafatos, E. Madueno, J. Modolell and D.M. Glover, 1998 P element insertion alleles of essential genes on the third chromosome of *Drosophila melanogaster*: Correlation of physical and cytogenetics maps in chromosomal region 86E-87F. Genetics **14**: 1697 - 1722). Using the technique of plasmid rescue, it has been possible to determine the nucleotide sequence surrounding the site of transposon insertion. The partial sequences of regions surrounding the P-element insertion site from distinct fly lines are

identified herein as SEQ ID Nos. 1-902. It is immediately evident to one skilled in the art how to use this information to clone a larger portion of nucleic acid containing the complete gene and thereafter express the encoded protein. Such techniques are disclosed for example in Sambrook et al (1989). It will be appreciated that the P-element may be inserted within a particular gene, or in regulatory sequences associated with a gene, such that P-element insertion affects expression of the gene, resulting in the lack of the expressed protein or expression of a dysfunctional form (ie. a protein with no activity or reduced activity such as a protein displaying less than 75%, 50%, 40%, 30%, 20%, 10%, 5% or 1% activity with respect to normal protein function). It is to be understood therefore that for the purposes of the present invention the term "essential gene" refers to the gene sequence itself as well as associated regulatory sequences which may be necessary for functional protein expression.

Each of the genes thus identified are shown to be potentially essential for cell or organism survival and/or reproduction based on the evidence of the effect of a complete loss of gene function caused by a mutation. Those skilled in the art will use a series of confirmatory steps to demonstrate that the transposon insertion site from which the identifiable gene sequence derives is indeed the site associated with lethality. Such confirmatory steps may comprise: repeated crossing to a wild-type strain (eg. Canton S) with investigation of the maintenance of a genetic linkage between the lethal locus and the transposon insertion site (cantonisation); reversion studies to show a complete correlation of lethal phenotype reversal and loss of the transposon; accurate genetic co-localisation of the lethal locus and the transposon insertion site using a set of deficiency (deletion) mutations covering the genomic region concerned.

Since the action of a protein encoded by such genes is thought to be essential for cell and/or organism survival or reproduction, it follows that chemical interference with the action or production of the protein will mimic the effect of mutation and result in the death of cells and the organism (or in the case of semi-lethality failure to reproduce or reduction in offspring). Thus, chemical interference with the action of each potentially essential gene and protein represents a way to kill cells and organisms, and specifically pests such as insects, arachnids, etc.

In a further aspect the present invention provides a polynucleotide fragment comprising nucleotides capable of encoding or partially encoding an essential gene for use in assays of the present invention. Typically the nucleotide fragment may be selected from any of the sequences identified by SEQ. ID. Nos. 1-902, or fragment thereof as described herein or other species homologue. More particularly the present invention provides means for obtaining essential proteins as encoded by the essential genes defined herein for use in assays of the present invention.

"Polynucleotide fragment" as used herein refers to a chain of nucleotides such as deoxyribose nucleic acid (DNA) and transcription products thereof, such as RNA. Naturally, the skilled addressee will appreciate the whole naturally occurring Drosophila or other "pest" species genome is not included in the definition of polynucleotide fragment.

The polynucleotide fragment can be isolated in the sense that it is substantially free of biological material with which the whole genome is normally associated *in vivo*. The isolated polynucleotide fragment may be cloned to provide a recombinant molecule comprising the polynucleotide fragment. Thus, "polynucleotide fragment includes double and single stranded DNA, RNA and polynucleotide sequences derived therefrom, for example,

subsequences of said fragment and which are of any desirable length. Where a nucleic acid is single stranded then both a given strand and a sequence or reverse complementary thereto is within the scope of the present invention.

In general, the term "expression product" refers to both transcription and translation products of said polynucleotide fragments. When the expression product is a "polypeptide" (i.e. a chain or sequence of amino acids displaying a biological activity substantially similar (eg. 98%, 95%, 90%, 80%, 75% activity) to the biological activity of an essential protein), it does not refer to a specific length of the product as such. Thus, the skilled addressee will appreciate that "polypeptide" encompasses *inter alia* peptides, polypeptides and proteins. The polypeptide if required, can be modified *in vivo* and *in vitro*, for example by glycosylation, amidation, carboxylation, phosphorylation and/or post-translational cleavage.

The present invention further provides an isolated polynucleotide fragment capable of specifically hybridising to a related polynucleotide sequence from another species. In this manner, the present invention provides probes and/or primers for use in *ex vivo* and/or *in situ* detection and expression studies. Typical detection studies include polymerase chain reaction (PCR) studies, hybridisation studies, or sequencing studies. In principle any specific polynucleotide sequence fragment from the identified sequences may be used in detection and/or expression studies. The skilled addressee understands that a specific fragment is a fragment of the sequence which is of sufficient length, generally greater than 10, 12, 14, 16 or 20 nucleotides in length, to bind specifically to the sequence, under conditions of high stringency, as defined herein, and not bind to unrelated sequences, that is sequences from elsewhere in the genome of the organism other than an allelic form of the sequence or non-

homologous sequences from other organisms.

"Capable of specifically hybridising" is taken to mean that said polynucleotide fragment preferably hybridises to a related or similar polynucleotide sequence in preference to unrelated or dissimilar polynucleotide sequences.

The invention includes polynucleotide sequence(s) which are capable of specifically hybridising to an essential polynucleotide sequence or to a part thereof without necessarily being completely complementary or reverse complementary to said related polynucleotide sequence or fragment thereof. For example, there may be at least 50%, or at least 75%, at least 90%, or at least 95% complementarity. Of course, in some cases the sequences may be exactly reverse complementary (100% reverse complementary) or nearly so (e.g. there may be less than 10, typically less than 5 mismatches). Thus, the present invention also provides anti-sense or complementary nucleotide sequence(s) which is/are capable of specifically hybridising to the disclosed polynucleotide sequence. If a specific polynucleotide is to be used as a primer in PCR and/or sequencing studies, the polynucleotide must be capable of hybridising to related nucleic acid and capable of initiating chain extension from 3' end of the polynucleotide, but not able to correctly initiate chain extension from unrelated sequences.

If a polynucleotide sequence of the present invention is to be used in hybridisation studies to obtain a related sequence from another organism the polynucleotide sequence should preferably remain hybridised to a sample polynucleotide under stringent conditions. If desired, either the test or sample polynucleotide may be immobilised. Generally the test polynucleotide sequence is at least 10, 14, 20 or at least 50 bases in length. It may be labelled by suitable techniques known in the art. Preferably the test polynucleotide sequence is at least 200 bases in length and may even be several kilobases in length. Thus, either a denatured sample or test sequence

can be first bound to a support. Hybridization can be effected at a temperature of between 50 and 70°C in double strength SSC (2xNaCl 17.5g/l and sodium citrate (SC) at 8.8g/l) buffered saline containing 0.1% sodium dodecyl sulphate (SDS). This can be followed by rinsing of the support at the same temperature but with a buffer having a reduced SSC concentration. Depending upon the degree of stringency required, and thus the degree of similarity of the sequences, such reduced concentration buffers are typically single strength SSC containing 0.1%SDS, half strength SSC containing 0.1%SDS and one tenth strength SSC containing 0.1%SDS. Sequences having the greatest degree of similarity are those the hybridisation of which is least affected by washing in buffers of reduced concentration. It is most preferred that the sample and inventive sequences are so similar that the hybridisation between them is substantially unaffected by washing or incubation in standard sodium citrate (0.1 x SSC) buffer containing 0.1%SDS.

Oligonucleotides may be designed to specifically hybridise to essential nucleic acid. They may be synthesised, by known techniques and used as primers in PCR or sequencing reactions or as probes in hybridisations designed to detect the presence of related material in a sample. The oligonucleotides may be labelled by suitable labels known in the art, such as, radioactive labels, chemiluminescent labels or fluorescent labels and the like. Thus, the present invention also provides oligonucleotide probes and primers for use in detecting essential genes from other organisms and which may be used in screening assays for pesticides.

The term "oligonucleotide" is not meant to indicate any particular length of sequence and encompasses nucleotides of preferably at least 10b (e.g. 10b to 1kb) in length, more preferably 12b-500b in length and most preferably 15b to 100b.

The oligonucleotides may be designed with respect to any of the sequences shown in SEQ ID Nos. 1-902 and may be manufactured according to known techniques. They may have substantial sequence identity (e.g. at least 50%, at least 75%, at least 90% or at least 95% sequence identity) with one of the strands shown therein or an RNA equivalent, or with a part of such a strand. Preferably such a part is at least 10, at least 30, at least 50 or at least 200 bases long. It may be an open reading frame (ORF) or a part thereof.

Oligonucleotides which are generally greater than 30 bases in length should preferably remain hybridised to a sample polynucleotide under one or more of the stringent conditions mentioned above. Oligonucleotides which are generally less than 30 bases in length should also preferably remain hybridised to a sample polynucleotide but under different conditions of high stringency. Typically the melting temperature of an oligonucleotide less than 30 bases may be calculated according to the formula of; 2°C for every A or T, plus 4°C for every G or C, minus 5°C. Hybridization may take place at or around the calculated melting temperature for any particular oligonucleotide, in 6 x SSC and 1% SDS. Non specifically hybridised oligonucleotides may then be removed by stringent washing, for example in 3 x SSC and 0.1% SDS at the same temperature. Only substantially similar matched sequences remain hybridised i.e. said oligonucleotide and corresponding test nucleic acid.

When oligonucleotides of generally less than 30 bases in length are used in sequencing and/or PCR studies, the melting temperature may be calculated in the same manner as described above. The oligonucleotide may then be allowed to anneal or hybridise at a temperature around the oligonucleotides calculated melting temperature. In the case of PCR studies the annealing temperature should be around the lower of the calculated melting temperatures for the two priming oligonucleotides. It is to be appreciated

that the conditions and melting temperature calculations are provided by way of example only and are not intended to be limiting. It is possible through the experience of the experimenter to vary the conditions of hybridisation and thus anneal/hybridise oligonucleotides at temperatures above their calculated melting temperature. Indeed this can be desirable in preventing so-called non-specific hybridisation from occurring.

It is possible when conducting PCR studies to predict an expected size or sizes of PCR product(s) obtainable using an appropriate combination of two or more oligonucleotides, based on where they would hybridise to the sequences SEQ ID Nos. 1-902. If, on conducting such a PCR on a sample of DNA, a fragment of the predicted size is obtained, then this is predictive that the DNA encodes a homologous sequence from a test organism.

The partial sequences of the essential genes, identified as SEQ ID Nos. 1-902, have been analysed in order to ascertain if there is any homology to previously sequences contained in nucleotide sequence databases such as the GENBANK and EMBL databases. Database searching has ascertained that a number of the nucleotide sequences show homology to sequences deposited in such databases. Some sequences show homology to sequences to which have been ascribed a function. However, the function of the gene/protein associated with the sequence may not have been suggested to be an essential gene/protein, the modulation of which may result in a lethal/semi-lethal phenotype. Other sequences show homology only to sequences for which no putative function has been ascribed. Finally, some sequences appear to show little or non-significant homology to sequences deposited in databases at the time of filing the priority application. Nevertheless, because of sequencing projects, such as the human and Drosophila genome sequencing projects many sequences have been deposited in such databases in the priority year. Thus, sequences which displayed little or no homology to

sequences in the databases on filing the priority application, may now show homology to sequences. However, the majority of such newly deposited sequences have no function ascribed to the sequence, or for that matter are known to relate or be associated with essential gene sequence.

Tables 1 and 2 summarise details of the sequences identified by the present invention. Those sequences which display homology to previously identified Drosophila gene sequences are identified under the class heading as GNL. Sequences which show a match with expressed sequence tags (ESTs) are represented under the class heading as EST. Sequences which show homology to genomic sequences are shoen under the class heading as GENO.

Additionally the class heading "IDNumber" is not of relevance to the present application; "Chr" relates to the chromosome from which the sequence was obtained; "feature" relates to the portion of sequence showing a match with the identified feature shown under "Name of match; and "AccNo" relates to the database accession number of the matching sequence.

Table 1 also refers to sequences identified by the present inventors (ie. SEQ ID Nos. 430-783 and 899-902) which do not show a clear match or homology to sequences in the database. Table 2 is essentially Table 1, updated, once information, such as the Drosophila genome sequencing project information had been submitted to a database. Most of the sequences which did not show a match or homology to sequences in the database, as shown in Table 1, as at the priority date, are shown in Table 2 as having homology to Drosphila sequences since this information became available during the priority year.

Additionally a number of sequences represented in Table 1 have now been found to relate to sequences which were previously published as being associated with sequences from genes known to be essential. These sequences are now shown in bold in Table 1.

Genes described herein generally fall into two classes:

A. Genes encoding proteins with recognizable similarity to proteins of known functional class e.g. protein kinases, neurotransmitter receptors; and

B. Genes encoding proteins of unknown function.

Among this set of genes are included genes with exact or a high degree of homology to known Drosophila Expressed Sequence Tags (ESTs) which provides evidence that they correspond to unknown genes that are expressed as messenger RNA. Other genes in this class have strong homology to ESTs from other organisms. All the remaining sequences are recognised as Drosophila genes on the basis of the genomic sequence.

The genes may be grouped on the basis of predicted related functions. This may be summarised as shown below, where the numbers correspond to SEQ ID Nos. disclosed herein:

1. GENES ENCODING RECEPTORS AND ASSOCIATED PROTEINS
SEQ ID Nos. 2. 3. 38, 157, 468, 471, 532, 538, 670 and 860.

2. GENES ENCODING CHANNELS
SEQ ID Nos. 39, 524, 798, 854, 864 and 896.

3. GENES ENCODING MEMBRANE PROTEINS (GENERAL; excluding receptors and channels
SEQ ID Nos. 6, 11, 13, 31, 199, 253, 299, 441, 448, 450, 559, 579, 581, 604, 717, 804, 813 and 856

4. GENES ENCODING KINASES including protein kinases
SEQ ID Nos. 62, 65, 66, 147, 185, 287, 332, 379, 393, 471, 486, 497, 547, 551, 583, 638, 685, 707 and 888.

5. GENES ENCODING PHOSPHATASES
SEQ ID Nos. 297, 420, 659 and 720.

6. GENES ENCODING GENERAL CELL SIGNALLING PROTEINS
SEQ ID Nos. 1, 22, 23, 69, 85, 102, 132, 159, 183, 229, 234, 313, 414, 421, 548, 625, 696, 715, 770, 771 and 885.

7. GENES ENCODING NTPases, including ATPases and GTPases
SEQ ID Nos. 11, 21, 24, 31, 40, 117, 227, 296, 356, 515 and 803.

8. GENES ENCODING INTRACELLUR STRUCTURAL, ORGANELLAR (except mitochondrial) AND SECRETORY SYSTEM PROTEINS
SEQ ID Nos. 29, 51, 55, 71, 195, 263, 301, 305, 333, 457, 560, 637, 743 and 753.

9. GENES ENCODING HEAT SHOCK PROTEINS
SEQ ID Nos. 41, 45, 49, 52, 64, 426 and 649.

10. GENES ENCODING DNA AND RNA ASSOCIATED PROTEINS (except transcriptional activators and repressors)
SEQ ID Nos. 5, 10, 20, 37, 72, 217, 274, 301, 303, 328, 364, 369, 394, 404, 466, 469, 516, 523, 549, 635, 636, 698, 725, 767, 802, 816, 818, 862, 893 and 898.

11. GENES ENCODING TRANSCRIPTION FACTORS, ACTIVATORS, REPRESSORS AND ASSOCIATED PROTEINS
SEQ ID Nos. 58, 61, 93, 411, 444, 451, 456, 492, 554, 633, 648, 684, 729, 831, 867 and 890.

12. GENES ENCODING PROTEIN SYNTHESIS, PROTEIN HANDLING AND ASSOCIATED PROTEINS
SEQ ID Nos. 15, 46, 47, 127, 243, 261, 321, 360, 367, 460, 558, 560, 759, 845 and 850.

13. GENES ENCODING PROTEIN MODIFICATION PROTEINS
SEQ ID Nos. 67, 244, 484, 776 and 873.

14. GENES ENCODING GENERAL PROTEIN DEGRADATION PROTEINS (except proteases)

SEQ ID Nos. 90, 116, 309, 326, 405, 611 and 765.

15. GENES ENCODING PROTEASES

SEQ ID Nos. 231, 346, 441, 485 and 742.

16. GENES ENCODING PROTEINS INVOLVED IN METABOLISM

SEQ ID Nos. 27, 34, 56, 83, 89, 108, 114, 121, 148, 188, 193, 196, 223, 230, 248, 278, 311, 312, 329, 353, 387, 416, 428, 435, 479, 521, 543, 564, 595, 597, 619, 631, 655, 677, 711, 716, 724, 726, 727, 738, 763, 773, 810, 825, 849, 862, 863, 866 and 892.

17. GENES ENCODING MITOCHONDRIAL AND ASSOCIATED PROTEINS

SEQ ID Nos. 12, 16, 32, 35, 40, 75, 118, 225, 514 and 842.

18. MISCELLANEOUS GENES CORRESPONDING OR HOMOLOGOUS TO KNOWN GENES OF DROSOPHILA OR OTHER ORGANISMS (other than the genes corresponding to the SEQ ID Nos listed for categories 1 to 17 above).

SEQ ID Nos. 4, 7, 8, 14, 18, 19, 26, 28, 42, 43, 44, 48, 54, 57, 60, 63, 68, 70, 73, 74, 78, 80, 87, 92, 95, 111, 134, 154, 161, 162, 168, 209, 213, 239, 246, 275, 276, 281, 295, 300, 304, 320, 322, 334, 335, 348, 358, 371, 372, 381, 389, 390, 410, 419, 442, 470, 499, 500, 566, 568, 621, 624, 632, 639, 654, 693, 695, 700, 709, 713, 730, 736, 745, 761, 766, 770, 790, 796, 829, 848, 853, 858, 863, 874, 881 and 901.

19. SEQ ID Nos corresponding to ESTs of Drosophila or other organisms for which no other information is available

SEQ ID Nos. 76. 77, 79, 81, 82, 84, 86, 88, 94, 96, 99, 101, 103, 104, 105, 106, 107, 110, 112, 113, 115, 119, 120, 122, 123, 124, 125, 126, 128, 129, 130, 131, 133, 135 to 146, 149 to 153, 160, 165, 166, 167, 169-181, 184, 186, 189

to 192, 194, 197, 198, 200, 201, 205 to 208, 211, 212, 214 to 216, 219 to 222, 224, 226, 228, 232, 235 to 238, 241, 242, 245, 247, 250 to 252, 254 to 260, 262, 264 to 273, 277, 280, 282 to 285, 288 to 294, 306 to 308, 310, 316 to 319, 323 to 325, 327, 330, 331, 336 to 345, 347, 349 to 352, 354, 355, 357, 359, 361 to 363, 365, 366, 368, 370, 373 to 376, 382 to 386, 388, 392, 395 to 403, 406 to 409, 412, 413, 415, 419, 422 to 424, 427, 429, 443 and 462.

20. ALL SEQ IDs NOT IN CATEGORIES 1 - 19 ABOVE CORRESPOND TO DROSOPHILA GENOMIC SEQUENCES WITHIN OR ADJACENT TO NEW GENES WITH NO CURRENT INFORMATION ON FUNCTION OR GENE EXPRESSION.

Assays for protein activities of known function are known in the art. Generally such assays are termed functional assays and may be conducted *in vitro* in a cell free or cell based system. A list of typical assays for some of the major classes of protein that are estimated to represent likely targets is exemplified herein. Where a functional assay is available, it is to be preferred to a ligand binding assay.

Assays for proteins of unknown function typically rely on assessment of ligand binding only, but other assays based on disturbance of chemical levels are well known to those of skill in the art.

The typical purpose of the assays described herein is to select for pesticides/insecticides, though in some cases lead compounds may have therapeutic activity, such as in inducing cell death which may be applicable in cancer therapy and other proliferative diseases. A relative specificity of action based on species groups or species may be achieved based on differences in protein sequence and structure, differences in protein expression, variations in development role and/or variations in degree of redundancy with related proteins.

The information disclosed herein teaches that the loss of the function of any of the proteins encoded by the genes comprising the partial sequences identified by SEQ ID Nos. 1-902, causes death of insects at some point during development, or causes severe physiological effects or reproductive failure. An insecticidal chemical compound will therefore be a compound that strongly modulates, either agonistically or antagonistically, the activity of such a protein. Thus, where the purpose of the assay is selection of insecticides, chemicals will be sought that interfere with the protein to modulate activity of the protein.

For proteins of known function with available functional assays, application of these assays will rapidly select a set of chemicals having the desired effect on the protein in the appropriate assay system.

In a second step, each member of the set of chemicals may then be tested directly for killing activity on insects. *Drosophila* itself is a convenient assay insect. In a typical fly killing assay, young flies are kept without fluid for a time, then transferred to vials containing filter paper dosed with a solution of the chemical to be tested. A range of chemical concentrations(eg. 10^{-2} - 10^{-10}M) may be used. After a defined treatment, flies are returned to normal conditions and observed. Rate of killing and percentage lethality are the parameters assessed.

In a third step, compounds with very effective killing activity on *Drosophila* may then be tested on pest species or accepted model pest related insects. Such pests include Dictyoptera (cockroaches); Isoptera (termites); Orthoptera (locusts, grasshoppers and crickets); Diptera (house flies, mosquito, tsetse fly, crane-flies and fruit flies); Hymenoptera (ants, wasps, bees, saw-flies, ichneumon flies and gall-wasps); Anoplura (biting and sucking lice); Siphonaptera (fleas); and Hemiptera (bugs and aphids), as well as arachnids such as Acari (ticks and mites). For

example, aphid species may be maintained on isolated lettuce plants: the time of death and the numbers of aphids falling dead onto paper traps beneath the plants after spraying with defined doses of the candidate chemical is assessed. As another example, lepidopteran pest larvae may be maintained on artificial media or plant leaves, which are treated with defined doses of chemicals, and survival is assessed.

The provision of candidate chemicals for use in the present invention are well known to those skilled in the art. For example libraries of compounds can be easily synthesised and tested. This is well described for example in: Applications of combinatorial technologies to drug discovery, 2. Combinatorial organic synthesis, library screening techniques, and future direction, J. Med. Chem.,1994,37,1385-1401.

For proteins of unknown function, the ligand binding assays outlined herein will also define a group of candidate chemicals. However, this group is likely to be large, since binding may occur to a number of different sites on the exposed surface of the protein, and binding alone does not predict the effect of ligand binding on the activity of the protein. Stringent selection among the candidate chemicals for those with the greatest affinity will define a set of chemicals small enough to be tested for insect killing. The use of *Drosophila* as a test organism enables large numbers of compounds to be assessed. Therefore the same procedure may be used as for proteins for which functional assays are available.

An alternative or additional procedure is to use a cellular killing assay as an intermediate step. For example, a gene of unknown function can be examined for location and timing of gene expression in tissues throughout development. The primary sites of tissue death may be determined by apoptosis assays or direct observation. In many cases, particular cell types e.g. nerve cells, can be defined as subject to death when the

protein is not expressed or inhibited. Appropriate cell types can be isolated from the appropriate tissue and developmental stage of *Drosophila* or a larger insect. Effects of candidate chemicals from the binding assay screen on survival of these cells in culture may then be ascertained, using commercially available live/dead cell assessment methods.

A further alternative or additional procedure is to express the protein target in a cell which has been manipulated genetically to contain a sensor for calcium ions, cyclic AMP or other components of cell signaling pathways. This may be achieved, for example, by generating transgenic *Drosophila* containing the gene encoding the protein with its expression driven by a promoter that is utilized in the cell type of choice. Alternatively, permanent cell lines of any suitable origin may be transfected, and lines expressing the protein permanently selected. In many cases, expression of an unknown protein will cause a shift in the level of cell signaling components, which will be detected by the sensor and can be read, for example, as a fluorescent or luminescent signal. The difference between the protein-expressing cells and control cells forms the basis of the assay. Effects of chemicals on the difference between protein expressing and control lines are assessed.

Proteins for all the assays described can be produced by cloning the gene for example into plasmid vectors that allow high expression in a system of choice e.g. insect cell culture, yeast, animal cells, bacteria such as *Escherichia coli*. To enable effective purification of the protein, a vector may be used that incorporates an epitope tag (or other "sticky" extension such as His6) onto the protein on synthesis. A number of such vectors and purification systems are commercially available.

The polynucleotide fragment can be molecularly cloned into a prokaryotic or eukaryotic expression vector using standard techniques and administered to a host. The expression vector is taken up by cells and the polynucleotide fragment of interest expressed, producing protein.

The cloning and expression of a recombinant essential polynucleotide fragment also facilitates in producing anti-essential antibodies and fragments thereof (particularly monoclonal antibodies).

It will be understood that for the particular polypeptides embraced herein, natural variations such as may occur due to polymorphisms, can exist between individuals or between members of the family. These variations may be demonstrated by (an) amino acid difference(s) in the overall sequence or by deletions, substitutions, insertions, inversions or additions of (an) amino acid(s) in said sequence. All such derivatives showing the recognised modulatory activity are included within the scope of the invention. For example, for the purpose of the present invention conservative replacements may be made between amino acids within the following groups:

(I) Alanine, serine, threonine;
(II) Glutamic acid and aspartic acid;
(III) Arginine and leucine;
(IV) Asparagine and glutamine;
(V) Isoleucine, leucine and valine;
(VI) Phenylalanine, tyrosine and tryptophan

Moreover, recombinant DNA technology may be used to prepare nucleic acid sequences encoding the various derivatives outlined above.

As is well known in the art, the degeneracy of the genetic code permits substitution of bases in a codon resulting in a different codon which is still capable of

coding for the same amino acid, e.g. the codon for amino acid glutamic acid is both GAT and GAA. Consequently, it is clear that for the expression of polypeptides from nucleotide sequences shown in SEQ. ID Nos. 1-902 or fragments thereof, use can be made of a derivative nucleic acid sequence with such an alternative codon composition different from the nucleic acid sequences shown in SEQ. ID Nos. 1-902.

The polynucleotide fragments of the present invention are preferably linked to regulatory control sequences. Such control sequences may comprise promoters, operators, inducers, enhancers, silencers, ribosome binding sites, terminators etc. Suitable control sequences for a given host may be selected by those of ordinary skill in the art.

A polynucleotide fragment according to the present invention can be ligated to various expression controlling sequences, resulting in a so called recombinant nucleic acid molecule. Thus, the present invention also includes an expression vector containing an expressible nucleic acid molecule. The recombinant nucleic acid molecule can then be used for the transformation of a suitable host. Such hybrid molecules are preferably derived from, for example, plasmids or from nucleic acid sequences present in bacteriophages or viruses and are termed vector molecules.

Specific vectors which can be used to clone nucleic acid sequences according to the invention are known in the art (e.g. Rodriguez, R.L. and Denhadt, D.T., Edit., Vectors: a survey of molecular cloning vectors and their uses, Butterworths, 1988, or Jones *et al.*,Vectors: Cloning Applications: Essential Techniques (Essential techniques series), John Wiley & Son. 1998).

The methods to be used for the construction of a recombinant nucleic acid molecule according to the invention are known to those of ordinary skill in the art and are inter alia set forth in Sambrook, et al. (Molecular Cloning: a laboratory manual Cold Spring Harbour Laboratory, 1989).

The present invention also relates to a transformed cell containing the polynucleotide fragment in an expressible form. "Transformation", as used herein, refers to the introduction of a heterologous polynucleotide fragment into a host cell. The method used may be any known in the art, for example, direct uptake, transfection transduction or electroporation (Current Protocols in Molecular Biology, 1995. John Wiley and Sons Inc.). The heterologous polynucleotide fragment may be maintained through autonomous replication or alternatively, may be integrated into the host genome. The recombinant nucleic acid molecules preferably are provided with appropriate control sequences compatible with the designated host which can regulate the expression of the inserted polynucleotide fragment, e.g. tetracycline responsive promoter, thymidine kinase promoter, SV-40 promoter and the like.

Suitable hosts for the expression of recombinant nucleic acid molecules may be prokaryotic or eukaryotic in origin. Hosts suitable for the expression of recombinant nucleic acid molecules may be selected from bacteria, yeast, insect cells and mammalian cells.

The present invention will now be described by way of non-limiting example and with reference to the attached sequence listing.

Example 1 - Generation of fly lines

The construct of the P{lacW} element used below is a defective P-element. A defective P-element is one which cannot transpose itself without the provision of a transposase enzyme from another source. Thus, once inserted into a site in the genome, a defective P-element will remain in position and will not distribute copies of itself. The reporter gene in P{lacW} is an *E. Col*: β-gal lacZ gene under the control of a weak promoter. This weak promoter, however, responds to enhancer elements in the neighbourhood of the insertion site to give a pattern of lacZ expression that is related, to a variable extent, to

the pattern of expression of the gene targeted. This provides temporal and/or tissue expression patterns which may be useful in deciding whether a gene/protein could be a potentially valuable target for insecticide or therapeutic development.

In addition to the reporter gene, P{lacW} carries a mini-*white* eye colour gene to identify flies that contain insertions. P{lacW} also contains a bacterial origin of replication and the β-lactamase gene coding for ampicillin resistance at its 3' end. This feature permits easy cloning of DNA flanking the insertion site of P{lacW} and further clone relevant genes (Bire, E., H. Vaessin, S. Shepherd, K. Lee, K. Mccall, S. Barbel, L. Ackermam, R. Carretto, T. Uemura, E. Grell, L.Y. Jan and Y.N. Jan, 1989 Searching for pattern and mutation in the Drosophila genome with a P-lacZ vector. Genes and Development **3**: 1273 - 1287).

The mutant flies, in which P{lacW} was inserted on the second chromosome are on the y w; P*{lacZ,w$^+$}*CyO genotype (Torok, T,G. Tick, M. Alvarado and I. Kiss, 1993 P-lacW Insertional Mutagenesis on the second chromosome of Drosophila melanogaster: Isolation of lethals with different overgrowth phenotypes. Genetics 135: 71 - 80). The mutant flies, in which P{lacW} was inserted on the third chromosome are of the *y w*; P{lacZ,w$^+$}TM3, *sb ser* genotype (Dèak, P., M.M. Omar, R.D.C. Saunders, M. Pal, O. Komonyi, J. Szidonya, P. Maroy, Y. Zhang, M. Ashburner, P. Benos, C. Savakis, I. Siden-Kiamos, C. Louis, V.N. Bolshakov, F.C. Kafatos, E. Madueno, J. Modolell and D.M. Glover, 1998 P element insertion alleles of essential genes on the third chromosome of *Drosophila melanogaster*: Correlation of physical and cytogenetics maps in chromosomal region 86E-87F. Genetics **14**: 1697 - 1722).

The genetic background of the w/w;P(lacW) mutants was equilibrated with that of the wild-type (Canton-S) strain by repeatedly backcrossing heterozygous w/w;P(lacW)/+ females (which carried the *w*$^+$ eye-color marker) to w(CS)

males for more than five generations. The *w*(CS) strain was derived by backcrossing w^{1118} flies to wild-type (Canton-S) flies for 10 generations; the w(isoCJ1) strain was derived from w(CS) and carries isogenic X, 2nd and 3rd chromosomes.

Example 2 - Plasmid Rescue and cDNA Cloning

Genomic sequences flanking the P-element were cloned by plasmid rescue using standard techniques (Drosophila: A practical Approach, the 2nd ed. 1998). Briefly, genomic DNA was digested with *Eco*RI, followed by ligation to form a rescue plasmid, which was propagated in *E. coli*. The rescue fragment then was ^{32}P-radiolabeled by random priming and used to screen plaques from a Drosophila genomic bacteriophage lambda library. The lambda genomic fragment was subcloned into the plasmid vector pBluescript, radiolabeled and used to probe a Drosophila adult head cDNA library and a Northern blot of adult whole fly polyA$^+$ RNA, etc.

Southern blotting was carried out essentially as described by Sambrook *et al.* (1989). Hybridization was carried out at 64°C in 6xSSC, 5xDenhardt's reagent, 0.5% SDS, 100μg/ml denatured, fragmented salmon sperm DNA. Filters were washed in 1xSSC and 0.1%SDS for 15 min, and then in 0.1XSSC and 0.1%SDS for 30 min.

Example 3 - DNA Sequencing

Prior to DNA sequencing, rescued plasmids were quantified by restriction digestion with EcoR1 to linearise followed by electrophoresis on a 1% agarose gel, comparisons being made to a Bacteriophage lambda 1kb marker ladder. For DNA sequencing 500ng-2μg of rescued plasmid was used in each sequencing reaction. Sequencing was carried out using a BigDye dideoxy terminator kit (Perkin-Elmer) with the following sequencing primers:-

1) 3' primer 5'-CGCACTTATTGCAAGCATACG-3' sequences into the rescued chromosomal DNA immediately 3' to the point of insertion (5' end of the chromsomal DNA insert)
2) 5' primer 5'-GCCACCTGACGTCTAAGAAACC-3' sequences the rescued chromosomal DNA from a point in the P-element vector 5' to the EcoR1 site, in a reverse orientation to Primer 1 (ie the 3' end of the chromosomal DNA)
NB. Sequence obtained from Primer 2 is only included in those sequences where the combined sequence runs yielded the complete insert of a particular clone.

The reactions were run on 5% polyacrylamide sequencing gels in 373A STRETCH PE Biosystems automated sequencer. Greater than 900 separate lethal/semi-lethal fly lines were identified by sequencing. The sequence obtained from these fly lines is represented in SEQ 10 Nos. 1-902.

DNA sequence analysis/storage was performed using Genejockey II and BLAST (Biosoft, Cambridge, UK)

Table 1 shows in summary details of the sequences obtained from the 902 distinct fly lines.

Table 2 shows in summary an updated version of Table 1 including references to sequences now contained in databases, but which were not disclosed until after the priority date of the present invention.

Example 4 - X-Gal Staining of Tissues

The procedure for X-gal staining of embryo is essentially as described by O'Kane (1998). Embryos are collected from yeasted apple/grape juice agar plates into a container with a nylon mesh screen at the bottom, dechorionated by dipping into 50% bleach for 4 minutes and washed thoroughly with water. Embryos are placed into an Eppendorf tube containing a mixture of 0.35ml fix solution (1% glutaraldehyde in PBS) and 0.7ml n-heptane and fixed for 15 minutes at room temperature on a rotating mixer. After removing heptane and fix solution from tube, embryos are washed three times for 10 min. in PBS and 0.1% Triton X-100, and resuspended in staining buffer with 0.2% X-gal

for 1-2 hours at 37°C. After staining, staining solution are removed and about 400μl of mixture solution (Glycerol : staining buffer = 2:1) are replaced. Embryos can then be mounted on a slide in a coverslip chamber.

For whole-mounts, larval, pupal and adult brains were dissected in PBS, and fixed in 4% paraformaldehyde for 20 min. They were then washed three times for 20 min in PBS, and stained with staining buffer and 2% X-gal for 1-2 h at 37°C (Ashburner, Drosophila. A Laboratory Manual, Plainview N.Y.: CSH Lab. Press 1989). They were then washed for 20 min in PBS, cleared overnight at 4°C with PBS/12.5% hydrogen peroxide, washed for 10 min with PBS, dehydrated through graded ethanol, and mounted in glycerol gelatin (Sigma).

To obtain sections, flies were mounted in "fly collars" (modified from Heisenberg and Böhl 1979), soaked in OCT embedding medium (Miles, USA) for 10 min and then embedded in the OCT medium. 12 μm serial sections of head or body were cut in a cryostat (Anglia Scientific) at -18°C. The sections were stained and mounted as described by Yang *et al.* (1995). Thereafter sections were examined and photographed on a Nomarski optical microscope.

Example 5 - *In situ* Hybridisation to Polytene Chromosome

In situ hybridisation to polytene chromosomes localises a DNA sequence (such as a gene, or am inserted P-element) on the physical DNA map of Drosophila, and may be related to the genetic map. For those insertion mutations which affect genes of known function, localisation of the P-element to the site where the gene mutated is known to reside is evidence that lethality does in fact result from insertion of the P-element in this gene.

Tagging the genes with a PlacW transposon allowed its immediate localization *in situ* to a precise cytological region using P element DNA as a probe. The procedure for *in situ* hybridization to third larval instar polytene chromosomes was essentially as described by Pardue (1986).

pBluescript (Strategene, USA) were labelled with Bio-16-dUTP by nick-translation. Hybridization was detected using 3,3' diaminobenzidine (DAB)/H_2O_2. After hybridization, the slides were stained with Giemsa and mounted using DPX which is a sliding mount commonly used by those skilled in the art.

As mentioned previously, the genes comprising any of the sequences disclosed herein may be cloned to allow expression of the associated protein and testing in an assay.

There now follows a non-exhaustive list of the types of assays which may be employed to test the ability of a candidate pesticide or therapeutic agent. In addition to the assays mentioned herein, the skilled addressee will be immediately aware of general texts such as Methods in Enzymology which is incorporated herein by reference. Additionally unless specifically mentioned the assays may be conducted at between 0°C-40°C, such as 4°C-35°C, or 20°C-30°C. Additionally the assays may be conducted at or around physiological pH.

1. Binding assays for proteins of unknown function

Ligands for any protein may be discovered by direct binding assays. In order to select true lead chemicals for insecticide or therapeutic development, these must be followed by insect killing assays or other functional assays as mentioned herein.

In binding assays one partner molecule is immobilized, and the other is labeled in some fashion (e.g. using a fluorescent tag, or by the incorporation of a radioactive isotope) and added free in solution. After incubation to allow molecular interaction, and a wash step, the amount of bound ligand is measured using an appropriate detection system. This may be used in a qualitative mode at first. Ligands showing significant binding may then be studied further by ensuring that the protein is in excess, and carrying out experiments with a dilution series of the

ligand at a set of known concentrations, typically from 10^{-2} - 10^{-10}M, such as 10^{-3} - 10^{-8}M

In the assay taught herein, the protein encoded by the essential gene is identified, and the chemical ligand is unknown. Therefore the protein may be purified using an affinity system and immobilized. The chemical ligands will be labeled, incubated with the immobilized protein, washed, and the amount of retained label assessed.

Proteins may conveniently be immobilized using an epitope or other affinity tags provided by the expression vector (see above), on a support material to which the appropriate antibody or binding agent for the tag is attached. The support material may be nitrocellulose membrane, Sephadex or other type of protein purification column support, or specialized beads such as those commercially available from Dynal or Promega. Alternatively, the protein may for example be biotinylated, and the same support materials derivatised with streptavidin (which has a very high affinity for biotin) used. Further, proteins may be modified chemically in a variety of ways, and covalently attached to support materials.

Nucleic acid or peptide ligands may conveniently be radioactively labeled by standard procedures. Organic chemical ligands may also be provided in radioactively labeled form. However, a more convenient labeling system for large scale screening by binding assays is the use of chemicals that are tagged with oligonucleotide sequence labels, or by other means. This allows many chemicals to be tested together initially, since each can be identified by the use of a PCR based detection system.

Monoclonal antibodies raised against a particular protein may be used to select chemicals that bind to particular regions of the protein - the epitope recognised by the antibody. In such an assay, chemicals are assessed for displacement or reduction in binding of the antibody. Remaining bound antibody is detected by a standard

fluorescently labeled second antibody.

2. Competitive Binding assays for proteins

For proteins belonging to families for which chemical or peptide ligands can be predicted, binding assays may be employed in a ligand-competition mode. This measures chemical interaction with the site on the protein at which the natural ligand binds, and is thus going to give a higher rate of significant hits. This type of assay is also more quantitative.

Examples of typical known ligands which would be labeled (typically radioactively) and used in displacement assays are: pharmacological agonists and antagonists, activators and inhibitors, neurotransmitters, growth factors and cytokines, cAMP, cGMP, enzyme cofactors such as NAD and FAD, regulatory polypeptides (e.g. calmodulin) and other subunits of multicomponent proteins.

A typical assay relies on the generation of purified protein as discussed above. In general, binding assays rely on labeled ligand, usually radiolabeled, to enable competition for the binding site to be detected. A set concentration (enough to saturate the binding site) of labeled ligand is incubated with a purified sample containing the purified protein. In a parallel tube, the test chemical is also added. Bound ligand/protein complexes are washed (to remove free ligand), precipated e.g. by TCA, collected with a cell harvester (for example) and the level of radioactivity measured. Displacement can be observed as a reduction in the amount of radioactivity detected in the assay. Enhancement of binding can also be observed in this type of assay, where radioactivity levels are increased - this indicates activity of the test chemical near to but not competing with the site of ligand interaction.

Examples of some possible functional assays

1. Kinases

Kinases are enzymes that transfer the terminal phosphate group of ATP and/or GTP to their substrate molecule. These enzymes have been shown to be involved in many cell processes including signal transduction, apoptosis and regulation of the cell cycle. Protein kinases are the largest known protein family and have been characterised in mammals, plants, fungi and microorganisms.

An assay of kinase activity generally requires two distinct steps: (1) transfer of the (labeled) terminal phosphoryl group of the nucleoside triphosphate donor to the substrate and (2) separation of the phosphorylated product from unutilized nucleotide. Step 1 is generally carried out in solution, with both the enzyme and the substrate in the liquid phase. Step 2 is usually accomplished by trichloroacetic acid (TCA) precipitation, by sodium dodecyl sulphate (SDS) gel electrophoresis, or by binding the labeled product to a solid support such as phosphocellulose paper or nitrocellulose membrane. These steps are then followed by detection of the amount of labeled phosphoryl that has been transferred to the substrate.

Step 1 can also be carried out with either the enzyme or the substrate immobilized on a solid support. For example, complex protein mixtures can be fractionated by SDS gel electrophoresis, blotted onto membrane, and then tested as potential substrates by incubating the membrane with a non-specific blocking agent followed by the desired protein sample plus labeled ATP.

Another variation on this type of assay involves detection of the phosphorylated form of a protein using a monoclonal antibody directed to the phosphorylated form. The amount of phosphorylation may then be assayed in an enzyme-linked immunosorbent assay (ELISA).

Typical assay of protein kinase activity

The kinase activity of a particular sample or protein can be assayed using histone H-1 (or other convenient protein kinase substrate) as a substrate to which the kinase transfers phosphate. For example in a reaction volume of 100μl containing 30 mM HEPES (pH 7.5), 5 μM $MgCl_2$, 40 μg of histone, 100μM $CaCl_2$, 10 μM [γ-32P] ATP and 1.25 mg/ml phosphatidylserine. Assays are started by the addition of 2.5m-units (arbitrary units, dilution series) of sample, incubated at 30°C for 10 minutes and terminated either by spotting on to P81 paper (Whatman) or by the addition of Laemmli buffer (Laemmli U.K., 1970 Nature 227, 680-685., or for more recent methods see Current Protocols in Molecular Biology Chapter 10, 1994-1997, eds Ausbel F.M. et al., Wiley). Spotting onto Whatman paper is followed by extensive washing in 75 mM orthophosphoric acid. The papers are then washed in ethanol, dried and incorporated radioactivity determined either by autoradiography, scintillation spectroscopy or phosphoimaging. After the addition of Laemmli sample buffer the sample is resolved on a 10% SDS-PAGE gel; the gel is dried and then autoradiographed. The amount of incorporated nucleotide is then determined using either autoradiography or phosphoimaging. It may be advantageous to add co-factors known to those skilled in the art that activate a particular kinase e.g. Calcium-dependent kinases would require calcium within the assay.

Such an assay is described for example in Wilkinson SE, Parker P and Nixon JS (1993) Biochem. J., **294**, 335-337. Further reference is made to Methods in Enzymology, 200:pp85-158,1991.

2. Protein Phosphatases

Protein phosphorylation provides one means of regulating cellular processes. Protein dephosphorylation by protein phophatases plays an equally important role. Phosphatases are involved in the removal of the phosphoryl group from proteins that have been phosphorylated by kinases. See for example Methods in Enzymology 201:pp389-468.

Typical assay for protein phosphatase activity

Assays for phosphatase activity can be carried out in the same way as a kinase assay. This would involve the pre-phosphorylation of for example histone by a kinase in the presence of radioactive ATP, followed by desphorylation by the test protein. The sample is then spotted onto P81 paper and the amount of radioactive ATP still incorporated is measured as previously described.

3. Adenylyl cyclases - cAMP formation from ATP

Measurement of intracellular cAMP.

Cyclic adenosine 3',5' monophosphate (cAMP) can be measured in tissue slices, dissociated tissue, cultured cells and membrane preparations.

Two procedures are currently used for measurement of cAMP: (1) radioimmunoassay and (2) the cAMP binding protein method

Radioimmunoassay uses antibody raised to acetylated cAMP and involves competition between cAMP in the sample and ^{125}I-labelled cAMP (Steiner, A.L., Wehmann, R.E., Parker, C.W. and Kipnis, D.M. (1972). *Adv. Cyc. Nucleotide Res.*, **2**, 51.). Following an overnight incubation, unbound cAMP is removed using charcoal. cAMP levels are quantified by comparison with a cAMP standard curve and expressed relative to protein content of sample. This method is sensitive in the femtomolar range if the sample cAMP and the standard

curve cAMP are acetylated before assay. Kits are available commercially (Amersham).

The cAMP binding protein method is based on competition between ^{3}H-labelled cAMP and sample cAMP for binding sites on the regulatory subunit of cAMP-dependent protein kinase (Gilman, A.G. (1970). *Proc. Natl. Acad. Sci. USA*, **67**, 305.). The procedure is analogous to radioimmunoassay but is more rapid because competition equilibrium is achieved in a 2 hour incubation. cAMP-dependent protein kinase preparation (Sigma) and binding protein assay kits (Amersham) are available commercially.

Enzyme immunoassay for cAMP. The Biotrack™ assay (Amersham Pharmacia Biotech) is an enzyme immunoassay in which the sample cAMP and peroxidase-linked cAMP compete for binding to antibody raised against acylated cAMP.

Measurement of adenylyl cyclase activity.

Adenylyl cyclase catalyses the formation of cAMP from ATP in the presence of Mg^{2+}. The main methods are: (1) the measurement of ^{32}P-labelled cAMP formed from $\alpha^{32}P$-labelled ATP and (2) the measurement of cAMP formed in a non-labelled reaction using either the radioimmunoassay or the binding protein assay.

Assay for adenylyl cyclase using $\alpha^{32}P$-labelled ATP: Radioactively labelled cAMP produced from $\alpha^{32}P$-ATP in an *in vitro* reaction is separated from unreacted substrate and radioactive contaminants by sequential chromatography steps on Dowex and alumina columns and measured by liquid scintillation counting (Salomon, Y., Londos, C., and Rodell, M. (1974). *Anal.Biochem.*, **58**, 541.). Crude or partially purified adenylate cyclase samples may contain contaminating activities that interfere with the assay. Problems with nucleoside triphosphatase are minimised using a high substrate concentration in the adenylate cyclase reaction and by including phosphoenol pyruvate and pyruvate kinase as

an ATP regenerating system. Degradation of ^{32}P-labelled cAMP can be prevented by including a high concentration of unlabelled cAMP in the reaction. The enzymatic reaction is terminated by addition of unlabelled ATP and by boiling for 2 minutes. Addition of [^{3}H]cAMP as a recovery label allows correction for differences in the performance of the individual chromatography columns. To isolate cAMP from the adenylate cyclase reactions the samples are first layered on a column of Dowex AG 50 WX 4 resin (200-400 mesh, H^{+} form) equilibrated in water. The cAMP has a greater affinity for the resin than ATP so the bulk of the [^{32}P] ATP can be washed of the column with water before eluting the cAMP directly onto an alumina column equilibrated with 0.1 M imidazole HCl, pH 7.5. The remaining [^{32}P] ATP binds to the alumina and the labelled cAMP is eluted using imidazole buffer. Samples are counted in ^{32}P and ^{3}H channels using a scintillation counter. Measurement of total [^{32}P] ATP and [^{3}H] cAMP allows calculation of pmols of cAMP present in the sample. Adenylate cyclase enzymatic activities are expressed as pmol cAMP formed per min per mg protein in the sample. The Dowex and alumina columns must be calibrated before use to determine elution profiles of ATP and cAMP but they may be regenerated after each assay and used repeatedly. The assay is sensitive, relatively simple and may be completed in one day. Apparatus for the double chromatography should be constructed from perspex to reduce risk from exposure to radioactivity.

Non-labelled adenylate cyclase reactions. Reactions contain ATP, Mg^{2+} and/or Mn^{2+} , an ATP regenerating system and an inhibitor of cAMP phosphodiesterase such as 3-isobutyl-1-methylxanthine (IBMX). Reactions are terminated by boiling and cAMP formed is measured by radioimmunoassay or cAMP protein binding assay.

See also Methods in Enzymology 195:pp3-21; and 288:pp326-339.

4. Guanylyl cyclases - cGMP formation from GTP

Measurement of guanylate cyclase activity.

Guanylate cyclase catalyses the hydrolysis of guanosine triphosphate (GTP) to cyclic guanosine 3',5' monophosphate (cGMP) in a reaction analogous to that of adenylate cyclase. Methodology used in the assay of guanylate cyclase activity is essentially the same as that for adenylate cyclase. Manganese is required as a cofactor for guanylate cyclase activity. Reactions are terminated by addition of HCl and boiling for 3 minutes.

Assay for guanylate cyclase using [^{32}P] GTP: This method depends on the separation of labelled-cGMP from unreacted substrate [^{32}P] GTP by sequential chromatography (Karczewski, P. and Krause, E.G. (1978). *Acta Biol. Med. Ger.*, **37**, 961.). Dowex 50 cation exchange columns and alumina columns are prepared and calibrated in exactly the same way as for the separation of cAMP except that the Dowex columns should be longer. cGTP is eluted from the alumina column with 0.2M ammonium formate buffer.

Non-labelled guanylate cyclase reactions. Reactions contain GTP, Mn^{2+}, a GTP regenerating system and IBMX. Reactions are terminated by boiling and cGMP formed is measured by radioimmunoassay using antibody against acetylated cGMP. Kits are available commercially (Amersham).

See also Methods in Enzymology, 195:pp345-354.

5. Phosphodiesterases - cAMP/cGMP hydrolysis

Assay of cyclic nucleotide phosphodiesterase activity.

Cyclic nucleotide phosphodiesterase catalyses the hydrolysis of the 3',5'-phosphodiester bond of the cyclic nucleotides, cAMP and cGMP.

The radioactive assay uses ^{3}H-labelled cAMP or cGMP and involves quantification of the reaction product (5'-nucleotide monophosphate) (Thompson,, W.J. and Appleman,

M.M. (1971). Biochemistry,10, 311.). The labelled nucleotide mono-phosphate (NMP) formed in the first reaction is converted to 5'-nucleotide in a second reaction by a 5'-nucleotidase present in snake venom (Alomone Labs, Jerusalem, Isreal). The labelled 5'-nucleotide is isolated by Dowex-1-chloride anion exchange chromatography and quantified by liquid scintillation counting.

See also Methods in Enzymology, 159:pp457-470; and 159:pp685-701.

6. ATPases - hydrolysis of ATP to ADP

Assay for adenosine 5'-triphosphatase

Adenosine 5'-triphosphatases (ATPases) catalyse the hydrolysis of ATP to ADP and inorganic phosphate in the presence of Mg^{2+}, Na^{+} and K^{+}. The colorimetric assay quantifies the inorganic phosphate released from ATP by measuring the A_{660}nm following treatment of the enzyme reaction with TCA and Taussky-Shorr Colour Reagent (Bonting, S.L., Simon, K.A., and Hawkins, N.M. (1961) *Arch. Biochem. Biophys.*, **95**, 416-423. Tausky, HH and Shorr, E. (1953) *J. Biol. Chem.*, **202**, 675-685.). Similar methods are used to assay guanosine 5'-triphosphatases (GTPases).

7. GTPases - hydrolysis of GTP to GDP

Assays essentially the same as for ATPases (see Sections 6 and 17). Commercial kits available.

8. Proteases

General Assay for Proteolytic activity Proteases

This assay is based on the proteolytic digestion of casein and the spectrophotometric detection of released aromatic amino-acids. Briefly, casein is incubated with the suspected protease and then acid precipitated. The solution is then filtered and the absorbance of the acid soluble phase is measured at 280-290nm. See for example W. Rick in

"Methoden der Enzymatichen Analyse", (H.U. Bergmeyer ed.) 3rd edition, 1046 and 1056. Verlag Chemie, Weinheim.

Example of a specific protease assay - Assay for the serine protease Chymotrypsin.

Endpoint titration with the fluorescent molecule 4-methyllumbelliferyl *p*-(N,N,N-triethylammonium) cinnamate. This compound is sensitive to 10^{-11} moles of enzyme with a 2 min reaction time, see for example G.W. Jameson, D.V. Roberts, R.W. Adams, W.S.A Kyle and D.T. Elmore. (1973) Biochem. J., 131, 107. See also Methods in Enzymology, 248:pp3-782.

9. Assays for secretion and import of proteins

These assays fall into three groups

A) Reconstitution in cell-free extracts

B) Reconstitution in semi-intact perforated cells

C) Assays for Endocytosis

A) Reconstitution in cell free extracts

The general principle of this type of assay is based on the detection of membrane fusion events and/or the delivery of protein contents using purified membrane compartments. The detection methods include immunodetection, fluorescence and release of chromogenic substances.

Example:- The detection of endocytic vesicle fusion *in vitro* using an assay based on the avidin-biotin association reaction.

The assay involves the use of two different populations of vesicles, each containing a different molecular probe conjugated to a marker protein. Upon fusion the probes bind to one another to generate a detectible signal, in this case the binding of avidin to biotin. Complexes are detected by an ELISA protocol (detecting the biotinylated protein e.g.

transferrin) and fluorescent detection of avidin conjugated ß-galactosidase, see for example William A. Braell in "Methods in Enzymology" **219**, 12-21 Academic Press inc. 1992.

B) Reconstitution using Semi-intact/perforated cells

Semi-intact/perforated cells are those which have lost a part of their plasma membrane by physical perforation. These assays can be done in Yeast or mammalian cells. Though lacking many soluble cytoplasmic factors, these cells retain their internal membrane and organellar structure and can efficiently reconstitute vesicular transport between compartments. They are also accessible to exogenously added factors such as antibodies and inhibitors.

Example:- Transport of a Protein between the Endoplasmic Reticulum and Golgi compartments.

This assay is based on the expression and transport of the Vesicular stomatitis virus (VSV) G protein. This viral glycoprotein has two Aspargine linked oligosaccharide chains which undergo extensive modifications as the protein transverses the ER and Golgi compartments. Oligosaccharide processing intermediates confer different electrophoretic mobilities on the VSV polypeptide, these intermediates can therefore be detected by SDS PAGE, see for example C.J.M. Beckers, D.S. Keller and W.E. Balch. Cell. **50**, 523 (1987).

C) Assays for Endocytosis

Assays for the endocytic pathway include those for detection of the binding of proteins to cell surface receptors, formation of clathryn coated endocytic vesicles, transport to the endosome, uncoating of the vesicles, delivery of the vesicle contents and recycling to the plasma membrane.

Example:- Detection of Functional Clathryn Coated Vesicles.

This assay involves the preparation of two vesicle fractions

i) The "donor" population containing ^{125}I-labelled transferrin.

ii) The "acceptor" vesicles, these being the clathryn coated vesicles under test. The acceptor vesicles contain internalised Anti-transferrin antibody.

The donor and acceptor populations are mixed in a solution containing cytosol and an ATP cocktail. Upon vesicle fusion a radiolabelled immunocomplex is formed. The vesicles are then solubilised and the mix passed through a *Staphlococcus aureus* column to isolate the immunocomplexes, which are then eluted from the column and the radioactivity measured, see for example P.G. Woodman and G. Warren in "Methods in Enzymolgy", **219**, 251 (1992)

10. Ribo/deoxyribo-nucleases - endo/exo-nuclease activity

Deoxyribonuclease

An endonuclease with preference for DNA. Pancreatic DNAse I yields di- and oligo-nucleotide 5' phosphates, pancreatic DNAse II yields 3' phosphates. In chromatin, the sensitivity of DNA to digestion by DNAse I depends on its state of organization, transcriptionally active genes being much more sensitive than inactive genes.

Ribonuclease

Widely distributed type of enzyme that cleaves RNA. May act as endonucleases or exonucleases depending upon the type of enzyme. Generally recognise target by tertiary structure rather than sequence. Ribonuclease E is an RNase involved in the formation of 5S ribosomal RNA from pre-rRNA. Ribonuclease F is stimulated by interferons and cleaves viral and host RNAs and thus inhibits protein synthesis. Ribonuclease H specifically cleaves an RNA base-paired to a complementary DNA strand. Ribonuclease P is an endonuclease

that generate t-RNAs from their precursor transcripts. Ribonuclease T is an endonuclease that removes the terminal AMP from the 3' CCA end of a non-aminoacylated tRNA. RNase T1 cleaves RNA specifically at guanosine residues. RNase III cleaves double-stranded regions of RNA molecules.

Endonuclease

One of a large group of enzymes that cleave nucleic acids at positions within the chain. Some act on both RNA and DNA (eg. S1 nuclease, EC.3.1.30.1, that is specific for single stranded molecules). Ribonucleases such as pancreatic, T1 etc. are specific for RNA, Deoxyribonucleases for DNA. Bacterial restriction endonucleases are crucial in recombinant DNA technology for their ability to cleave double-stranded DNA at highly specific sites.

Nuclease

An enzyme capable of cleaving the phosphodiester bonds between nucleotide subunits of nucleic acids.

Restriction Endonuclease

Class of bacterial enzymes that cut DNA at specific sites. In bacteria their function is to destroy foreign DNA, such as that of bacteriophages (host DNA is specifically modified at these sites). Type I restriction endonucleases occur as a complex with the methylase and a polypeptide that binds to the recognition site on DNA.. Type II restriction endonucleases are the classic experimental tools. They have very specific recognition and cutting sites. The recognition sites are short, 4-8 nucleotides, and are usually palindromic sequences. Because both strands have the same sequence running in opposite directions the enzymes make double-stranded breaks, which, if the site of cleavage is off-centre, generates fragments with short single-stranded tails; these can hybridise to the tails of

other fragments and are called sticky ends. They are generally named according to the bacterium from which they were isolated (first letter of genus name and the first two letters of the specific name). The bacterial strain is identified next and multiple enzymes are given Roman numerals. For example the two enzymes isolated from the R strain of *E. coli* are designated *Eco* RI and *Eco* RII. The more commonly used restriction endonucleases are known to those skilled in the art, but may be found in manufacturers catalogues, such as New England Biolabs, USA.

Ref: Definitions taken from the Dictionary of Cell Biology (Second Edition), Academic Press.

General Assay

All of the above nucleases cleave DNA and/or RNA, therefore a general assay would be to incubate unknown/test protein/chemical with a known quantity and type of DNA or RNA for a given time, and separate the products using gel electrophoresis along with a known set of standards. Any nuclease activity will be readily visible on the gel. Once nuclease activity has been detected, direct comparisons can be made with the DNA cleavage patterns generated by known nucleases in order to identify the type of nuclease involved.

11. DNA metabolism - ligase, topoisomerase, etc

DNA glycosidase

Class of enzymes involved in DNA repair. They recognise altered bases in DNA and catalyse their removal by cleaving the glycosidic bond between the base and the deoxyribose sugar. At least 20 such enzymes occur in cells.

DNA ligase

Enzyme involved in DNA replication. The DNA ligase of *E.coli* seals nicks in one strand of double-stranded DNA, a reaction required for linking precursor fragments during discontinuous synthesis on the lagging strand. Nicks are breaks in the phosphodiester linkage that leave a free 3'-OH and 5'-phosphate. The ligase from phage T4 has the additional property of joining two DNA molecules having completely base-paired ends. DNA ligases are crucial in joining DNA molecules and preparing radioactive probes (by nick translation) in recombinant DNA technology.

DNA methylation

Process by which methyl groups are added to certain nucleotides in genomic DNA. This affects gene expression, as methylated DNA is not easily transcribed. The degree of methylation is passed on to daughter strands at mitosis by maintenance DNA methylases. Accordingly, DNA methylation is thought to play an important developmental role in sequentially restricting the transcribable genes available to distinct cell lineages. In bacteria, methylation plays an important role in the restriction systems, as restriction enzymes cannot cut sequences with certain specific methylations.

DNA/RNA synthesis

DNA polymerase and RNA polymerase are enzymes involved in template-directed synthesis of DNA from deoxyribonucleotide triphosphates and RNA from ribonucleotide triphosphates.

Repair nuclease

Class of enzymes involved in DNA repair. It includes endonucleases that recognise a site of damage or an incorrect base pairing and cut it out, and exonucleases that

remove neighbouring nucleotides on one strand. These are then replaced by a DNA polymerase.

Topoisomerase

An enzyme capable of altering the degree of supercoiling of double-stranded DNA molecules. Various topoisomerases can increase or relax supercoiling, convert single-stranded rings to intertwined double-stranded rings, tie and untie knots in single stranded and duplex rings, catenate and decatenate duplex rings. Topoisomerase II of *E.coli* is commonly known as gyrase.

General Assay

All of the above act to modify the structure of DNA. For each enzyme involved in DNA metabolism, a corresponding assay is available commercially.

12. Transcription factors

Transcription Factor Assays

Transcription factor activity lies in the centre of a signalling cascade that begins at the cell surface by the activation of a receptor. Intracellular signal transduction events activate or repress specific transcription factors, which in turn regulate the expression of specific genes.

The activity of a transcription factor can be assessed by linking the appropriate regulatory sequence to a reporter gene encoding among other reporters β-galactosidase, Chloramphenicol acetyl transferase (CAT), luciferase and green fluorescent protein (GFP) in an engineered plasmid vector. This vector is used to transfect a cell line and the activity of the transcription factor of interest analysed by measuring the amount of reporter activity (Brannon, M. et al (1997) Gen. Dev. 11, 2359.).

Of the many different strategies available for using genetic reporters, luciferase offers the most ideal situation because the reporter measurements are nearly instantaneous, exceptionally sensitive and there is little or no endogenous activity in the host cells to interfere with quantitation. Firefly luciferase (Ow, D et al (1986) Science 234, 856.) is by far the most commonly used of bioluminescent reporters. The enzyme catalyses a two-step oxidation reaction to yield light at 550-570nm that can be detected by the use of a luminometer. The assay can be adapted for use with single or multiple samples depending on the type of luminometer available, i.e. tube or plate.

The above is an *in vivo* transcription factor assay requiring the transfection of an appropriate cell line with the reporter vector. However, an *in vitro* method for transcription/DNA binding factor analysis also exists.

The gel shift or electrophoretic mobility shift assay provides a simple and rapid method for detecting sequence-specific binding proteins, such as transcription factors (Ausubel, F.M. et al. (1989) In: Current Protocols in Molecular Biology, Vol. 2, John Wiley and Sons, New York.). The assay is based upon the observation that complexes of protein and DNA migrate through a non-denaturing polyacrylamide gel more slowly than free DNA fragments or double stranded oligonucleotides. The gel shift assay is performed by incubating a purified protein or a complex mixture of proteins such as a nuclear extract preparation with a ^{32}P labelled DNA fragment containing the putative binding site. The reaction products are then analysed on a nondenaturing polyacrylamide gel. The specificity of the DNA-binding protein for the putative binding site is established by competition experiments using DNA fragments or oligonucleotides containing a binding site for the protein of interest.

13. Apoptosis

Apoptosis is the physiological process by which unwanted or useless cells are eliminated during development and other normal biological processes, e.g. tissue homeostasis, embryogenesis, induction and maintenance of immune tolerance, development of the nervous system and endocrine-dependent tissue atrophy.

The process of apoptosis involves a cascade of cytoplasmic and nuclear events that result in a series of morphological changes and eventually cause the demise of the cell. Apoptosis can be initiated by a variety of different stimuli that lead to a convergence of biochemical signalling pathways into a common collection of executioner molecules.

In the early stages of apoptosis, changes occur at the cell surface and plasma membrane. One of these plasma membrane alterations is the translocation of phosphatidylserine (PS) from the inner side of the plasma membrane to the outer layer, where PS becomes exposed at the external surface of the cell. Mitochondrial physiology is disrupted in cells undergoing apoptosis. Permeability is altered and specific protease activators are released. Specifically, the discontinuity of the outer mitochondrial membrane results in the redistribution of cytochrome C to the cytosol followed by subsequent depolarisation of the inner mitochondrial membrane. Cytochrome C release further promotes apoptosis by the activation of the caspases, cysteine proteases. Active caspases participate in a cascade of cleavage events, which disable key homeostatic and repair enzymes and bring about a systematic structural disassembly of dying cells. The biological substrates of caspases include poly(ADP ribose) polymerase (PARP), DNA-dependent protein kinase (DNA-PK), lamins, topoisomerases, Gas2, protein kinase C (PKC), sterol regulatory element binding proteins (SREBP), U1-70kDa protein and Huntingtin protein.

The biochemical hallmark of apoptosis is the fragmentation of genomic DNA, an irreversible event that commits the cell to die and occurs before changes in plasma membrane permeability.

In concert with increased understanding of the physiological events that occur during apoptosis, a number of assay methods have been developed for its detection. These assay methods can measure one of the following apoptotic parameters:

1. Fragmentation of DNA in populations of cells or in individual cells, in which apoptotic DNA breaks into different length pieces.

2. Alterations in membrane asymmetry. PS translocates from the cytoplasmic to the extracellular side of the cell membrane.

3. Activation of apoptotic caspases.

4. Release of cytochrome C into the cytoplasm by mitochondria.

Each provides the researcher with a different level of information as each of these events occurs at a different stage in apoptosis.

The early proteolytic events of apoptosis can be monitored using an adaptation of the absorbance-based assay originally devised by Thornberry, N.A. (1994) Interleukin-1 beta converting enzyme. Meth.Enzymol. 244, 615. The colorimetric substrate (Ac-DEVD-pNA) is labelled with the chromophore p-nitroaniline(pNA). pNA is released from the substrate upon cleavage by DEVDase. Free pNA produces a yellow colour that is monitored by a photometer at 405nm. The amount of yellow colour that is produced upon cleavage is proportional to the amount of DEVDase activity present in the sample. The potent, irreversible, pan-caspase inhibitor benzoxycarbonyl-val-ala-asp fluromethyl ketone (Z-VAD-FMK), (Zhou Q, Krebs JF, Snipas SJ, Price A, Alnemri ES, Tomaselli KJ, Salvesen GS Biochem 37 10757 (1998) can be used as a negative control and it is suggested that apoptosis be

induced by the addition of Fas or TNF agonist antibodies.

The protocol can be used to test multiple samples by performing the assay in a total volume of 100ml using cells cultured in 96 well plates. The absorbance produced by each sample is read using a plate reader.

See also general reference Methods in Enzymology 322 pp3-522.

14. Calcium

Calcium dynamics

In a multicellular organism, cell communication is essential to regulate the different activities of specialised tissues. In all animal cells, there are conserved intracellular second messenger pathways. For many of these, calcium is an important second messenger. In nerve cells, muscle and other cells, modulation of intracellular calcium activity from typical resting levels of 100 nM regulates many short and long-term processes. Measurement of calcium can thus be of great utility in following the responses of transgene products to applied pharacological agents including insecticides.

Calcium dynamics may be detected directly or indirectly by a range of methods; including but not restricted to: a) transgenic apoaequorin, a calcium-sensitive luminescent protein; b) other methods that monitor intracellular calcium concentration; c) other methods that monitor the operation of intracellular calcium signalling pathways; d) methods that monitor the operation of other types of signalling pathway; e) methods that monitor neuronal electrical potentials.

For example, transgenic apoaeqorin has been used to monitor calcium dynamics in the intact Drosophila renal system and the intact Drosophila brain (Rosay et al (1997) J. Cell. Sci. 110, 1683-1692; O'Donnell et al. (1998) Am. J. Physiol. 43(4), R1039-R1049.). It has also been used to provide a bioluminescent assay for agonist activity against

G protein coupled receptors (Stables et al. (1997) Anal. Biochem. 252, 115-126). Transgenic apoaequorin can thus be used to assess the effect of an exogenous gene on intracellular calcium dynamics, the method comprising detecting a pattern of calcium dymamics in cells, tissues or organisms expressing the exogenous gene, and comparing said pattern with a pattern of calcium dynamics in cells, tissues or organisms without said exogenous gene.

Additionally, fluorescent probes (such as fura-2, indo-1, quin-2) show a spectral response upon binding calcium and it is then possible to detect changes in intracellular free calcium concentrations using fluorescence microscopy, flow cytometry and fluorescence spectroscopy. Most of these fluorescent indicators are variations of the nonfluorescent calcium chelators EGTA and BAPTA (Cobbold and Rink (1987) Biochem. J., **248**, 313.).

New fluorescent indicators for calcium called "cameleons" may also be used and are genetically encoded without cofactors and are targetable to specific intracellular locations. These so-called "cameleons" consist of tandem fusions of a blue- or cyan-emitting mutant of the green fluorescent protein (GFP), calmodulin, the calmodulin-binding peptide M13, and an enhanced green- or yellow-emitting GFP. Binding of calcium makes calmodulin wrap around the M13 domain, increasing (Miyawaki et al., (1997) Nature, **388**, 882-887.) or decreasing (Romoser et al., (1997) JBC, **272**, 13270-13274.) the fluorescence resonance energy transfer between flanking GFPs.

Additionally, potentiometric optical probes may be used. Potentiometric optical probes measure membrane potential in organelles and in cells. In conjunction with imaging techniques, these probes can be employed to map variations in membrane potential along neurons and among cell populations with high spatial resolution and sampling frequency (Rohr and Salzberg (1994) Biophys. J., **67**, 1301.).

Additionally, GFP-based reporter genes that monitor intracellular cAMP dynamics may be used, and to monitor intracellular pH changes (Miesenbock et al. (1998) Nature 394, 192-5).

15. cAMP

Effects on dynamics of intracellular cAMP as reported by appropriate dyes or reporter constructs (eg. aequorin). See section 13.

16. Voltage

Analysis of transmembrane potential permits study of the elements which mediate electrical behaviour of cells. This form of study may be undertaken in a number of ways, including: voltage (patch) - clamping and the use of voltage sensitive dyes.

Patch clamping

In brief, this involves sealing a blunt micropipette tip to a cell membrane. This is termed a **gigaseal**. The gigaseal electrically isolates the whole cell or a patch of the membrane allowing detection of picoampere, ionic currents while accurately controlling the voltage. This form of analysis may be utilised in the study of cultured cells, tissue slices or recombinant ion channels expressed post DNA transfection in heterologous cells. Whole cell recording measures the activity of the full complement of active channels in a cell; typically specific populations of channels are isolated using channel-blocking agents. It is also possible to isolate single ion channels, providing information on the unitary conductance and kinetic behaviour of individual channels, and allow the factors which alter these properties to be studied in exquisite detail (Crawley et al., 1997 Neurophysiology Current protocols in neuroscience Volume 1 [John Wiley and Sons,

Inc.]). Patch Clamp techniques are widely used and cited throughout scientific literature (Siegel M.S and Isacoff E.Y (1997) Neuran 19, 735 - 741; Sensi S., Canzoniero L.M, Yu S.P, Ying H.S, Koh J.Y, Kerchner G.A, Choi D.W. (1997) J. Neuroscience 17, 9554 - 9564; Piller S.C, Jans P., Gage P.W, Jans D.A (1998) Proc. Natl. Acad. Sci. USA, 95, 4595 - 4600; Maric D., Maric I., Wen X., Fritschy J.M, Sieghart W., Barker J.L, Serafini R., (1999) J. Neuroscience 19, 4921 - 4937). An example of how this kind of analysis may be used is outlined below.

Whole cell patch clamp recording to study the effects of a viral protein on whole cell currents of cultured hippocampal neurons.

Whole cell currents represent the integrated channel activity over the whole cell. Cultured cells on coverslips were perfused with bath solution (140 mM NaCl, 5 mM KCl, 3 mM $CaCl_2$, 2 mM $MgCl_2$, 10 mM glucose, 10 mM TES [pH7.3]) at room temperature (23 to 28C). Pipettes made from boroscilate glass were fire polished and filled with pipette solution normally containing 150 mM NaCl, 0.5 mM $CaCl_2$, 2 mM $MgCl_2$, 5 mM EGTA and 10 mM TES [pH7.3]. Reversal potentials were determined experimentally by altering the holding potential until currents reversed direction and the potential for zero current was recorded. Cells were routinely clamped at - 60 mV. Whole cell currents were recorded both before and after the addition of purified viral protein by using an Axopatch 200A. Viral protein @ 0.6 nM in bath solution was applied directly onto patched cells through gravity fed drug delivery tubing. whole-cell currents were filtered at 5 or 10 kHz, digitized at 44 kHz, and stored on videotape. For data analysis currents were replayed through the same system and digitized using an A to D converter interfaced with an IBM-compatible computer. Inward currents are depicted as downward deflections from the zero current level.

Electrical potential may also be measured using voltage sensitive dyes. e.g. Oxonol VI / Bis-oxonol (Dall'Asta V., Gahi R., Orlandini G., Rossi P.A, Rotoli B.M, Sala R., Bussolati O., Gazzola G.C, (1997) Experimental Cell Research 231, 260 - 268; Salvador J.M, Inesi G, Rigaud J.L, Mata A.M, (1998) J. Biol. Chem. 273, 18230 - 18234).

In the study by Salvador et al., 1998, transmembrane electrical potential was measured by analysis of the differential absorption (625 - 603 nm) of 2 μM oxonol using dual wavelength spectrophotometry. The medium comprised Pipes buffer, pH7.1; 0.42 μg/ml calmodulin; 5mM MgCl2 and 2 μM Oxonol VI. The callibration was performed by several additions 130 mM KCl to the medium in the presence of 1 μM of the K^+ ionophore valinomycin and in the absence of ATP. Absorption changes were standardized using the Nernst equation. Within a range 0 - 40 mV Absorption by oxonol demonstrates a linear increase with increasing membrane potential. This proportionality permits straightforward assay of changes in membrane potential.

Dall'Asta et al., 1997 visualize changes in membrane potential using Bis-oxonol. Bis-oxonol is a fluorescent dye which distributes across biological membranes according to the membrane potential and binds to hydrophobic components: since the quantum yield of the dye increases with binding, the fluorescence of the cells incubated in a medium containing the dye increases with depolarization and decreases with hyperpolarization.

17. Receptors/ion channels

Ion Channels/Receptors

Neuronal signaling depends on rapid changes in the electrical potential difference across nerve cell membranes. These rapid changes in potential are made possible by ion channels, a class of integral membrane proteins that traverse the cell membrane. These channels have three important properties: (1) they conductions, (2)

they recognise and select among specific ions, and (3) they open and close in response to specific electrical, mechanical, or chemical signals. [Principles of Neural Science, (Kandel and Schwartz), Chapter 5 Ion Channels]. Ion channels are large integral membrane glycoproteins, which have a central aqueous pore that spans the entire width of the membrane. Many ion channels are made up of two or more subunits, which may be identical or distinct. Three major signals gate ion channels: voltage (voltage-gated channels), chemical transmitters (transmitter-gated channels), and pressure or stretch (mechanically-gated channels). Gating involves a conformational change of the channel in response to the above stimuli.

Several major classes of ion channels have now been identified. Primary sequence information has been used to suggest the structure of different channel proteins. Efforts to determine secondary structure rely on X-ray crystalography. However, additional information can be obtained by comparing the primary amino acid sequence of related channels from different species and identifying regions of sequence homology, suggesting the importance of such regions in channel structure and function. Further insight into structure-function relationships can be obtained from sequence homologies among different, but related, channels. Such homologous regions are likely to underlie a common biophysical function shared by the different channels, i.e. Voltage-gated versus transmitter-gated channels.

The flux of ions through ion channels is passive, requiring no expenditure of metabolic energy. The direction and eventual equilibrium for this flux is determined not by the channel itself, but rather by the electrochemical driving force across the membrane. Ion channels select the type of ions that they allow to cross the membrane through physio-chemical interaction between the ion and various amino acid residues that line the walls of the channel pore (on the basis of ionic charge),

allowing either cations or anions to permeate. Some cation-selective channel types are relatively non-selective, passing Na^+, K^+, Ca^{2+}, and Mg^{2+}. However, most cation-selective channels are more selective; each one is permeable primarily to a single type of ion, either Na^+, K^+, or Ca^{2+}. All known types of anion-selective channel are permeable to Cl^-. Note that the $Ca2^+$ influx controlled by channels can alter many metabloic processes within cells, leading to activation of various enzymes and proteins. $Ca2^+$ influx also acts as a trigger for the release of neurotransmitter.

The activity of channels can be modified by cellular metaboloic reactions, including protein phosphorylation, by various channel blockers, toxins, poisons, and drugs. Channels are important targets in various diseases, eg myasthenia gravis and cystic fibrosis.

Molecular analysis: starting with an unknown chemical for which no information is available, and depending on the size of the starting molecule, peptide sequence can be obtained either directly or by using the chemical bound to a column to purify the target molecule in the cell (e.g., benzodiazepine affinity chromatography purification columns were used to isolate and identify the first cDNA clones encoding GABA receptor subunits in 1987 - Schofield P.R, Darlison M.G, Fujita N., Burt D.R., Stephenson F.A, Rodriguez H., Rhee L.M, Ramachandran J., Reale V., Glencorse T.A, Seeburg P. And Barnard E.A (1987) Nature 328, 221 - 227). From peptide sequence, it is possible, by back-translation, to identify the nucleotide sequences from which the peptide may be translated, and based on the codon usage of a particular organism, best-guess oligonucleotides can be synthesized and used to screen species-specific cDNA libraries. Any cDNA clones identified can then be sub-cloned, sequenced and the primary sequence analysed for known sequence homologies with BLAST database searches. The full-length sequence, cDNA and corresponding expressed

protein, can then be subjected to standard biochemical and molecular characterisation procedures.

Functional analysis: single-channel recording can measure the activity of a single protein molecule (electrophysiology). The patch clamp technique has made it possible to measure directly the activity of single ion channel molecules by recording the unit current flow through single open channels. Expression of cRNAs in the Xenopus oocyte system, cDNA in transfected cell lines or whole tissue slice cultures can be used.

Example of cRNA expression in Xenopus oocyte: pure mRNA, produced by *in vitro* transcription from cDNA, is microinjected into Xenopus oocytes and pulses of known compounds (eg gamma amino butyric acid (GABA), glutamate, etc) can be superfused over the oocyte while recording membrane currents under voltage-clamp conditions. Current response to applied compound can be measured.

Two main gene families:

1. Voltage gated channels:	Na^+, K^+ and Ca^{2+}
2. Ligand-gated ion-channel receptors: either Integral (pore forming, Ion gating) or Second messenger systems (associated)	cation:(nACh), 5HT, glutamate, anion: GABA Glycine

Methods

1. cloning: sequence analysis primary and secondary
2. structure: crystalography, in situ, immunocyto, immunohisto, immunoEM
3. function: electrophysiology: slice culture/patch clamp transfection/patch clamp

Voltage-Gated Channels

In nerve cells at rest (membrane potential: -65mV), the steady Na^+ influx through non-gated channels is balanced by steady K^+ efflux, so that the membrane potential is constant. This steady state balance changes when the cell is sufficiently depolarised to trigger an action potential. A transient depolarising potential, such as excitatory synaptic potential, causes some voltage-gated Na^+ channels to open, and the resultant increase in membrane Na^+ permeability allows Na^+ influx to outstrip the K^+ efflux. Thus, a net influx of positive charge flows through the membrane, and positive charges accumulate inside the cell, causing further depolarisation. The increase in depolarisation causes more voltage-gated Na^+ channels to open, resulting in a greater influx of positive charge, which accelerates the depolarisation further.

This regenerative, positive feedback cycle develops explosively, driving the membrane potential toward the Na^+ equilibrium potential of +55mV. Because K^+ efflux continues through the K^+ channels, the membrane potential never actually reaches the equilibrium potential of sodium. A slight diffusion of Cl^- into the cell also counteracts the depolarising tendency of the Na^+ influx.

As depolarisation continues, it slowly turns off, or inactivates, the voltage-gated Na^+ channels. That is, the Na^+ channels have two types of gating mechanisms: activation, which rapidly opens the channel in response to depolarisation, and inactivation, which slowly closes the channel if depolarisation is maintained. The second repolarising process results from the delayed opening of voltage-gated K^+ channels. The delayed increase in K^+ efflux combines with a decreased Na^+ influx to produce a net efflux of positive charge from the cell, which continues until the cell has repolarised to its resting membrane potential.

Intracellular recording: this technique uses two glass capilliary electrodes full of an ionic conductor solution (usually 3M KCl). To measure the resting membrane potential, an intracellular electrode is inserted into the nerve cell (grown in culture or via slice culture) - the pipette acts as a salt bridge, providing electrical connection between the cytoplasm and a metal electrode that is connected to the electronic apparatus. The second extracellular electrode can be used to confirm resting potential and/or stimulate the cell. Both electrodes are connected to a voltage amplifier, which in turn is connected to an oscilloscope that displays the amplitude of the membrane potential (-65mV at rest).

Ligand-Gated Channels - Integral Channel

(e.g., nicotinic acetylcholine (nACh) receptor, 5 hydroxytryptamine (serotonin) ($5HT_3$) receptor, glutamate, $GABA_A$, Glycine)

A transmembrane ion channel whose permeability is increased by the binding of a specific ligand, typically a neurotransmitter at a chemical synapse. The permeability change is often drastic; such channels let through effectively no ions when shut, but allow passage at up to 10^7 ions s^{-1} when a ligand is bound. These receptors have been found to share considerable sequence homology, implying that there may be a family of structurally related ligand-gated ion channels.

Ion channel receptors are composed of 4 or 5 subunits, which may be the same or different, each of which contains 4 or 5 membrane-spanning α-helical regions. These α-helices are thought to align to form the pore of the channel, through which ions can flow. The characteristics of each channel is determined by the type of subunits that are present in each receptor subtype. Annals of the New York Academy of Sciences (1999) Volume 868) Current flow depends on the number of open channels, the concentration

of the transmitter, channel conductance and membrane potential.

Receptor specific assays will have to be created for each receptor/ion channel under investigation. The best/easiest way to do this is to create permanent cell lines expressing a particular combination of receptor subunits in order to form particular receptor subtypes. There are many examples of these in the literature, and of the differences in receptor characteristics when different combinations of receptor subtypes are expressed. Initial assays established by the inventors will focus on the most clinically relevant subtype(s) of each receptor. With these permanent cell lines, functional assays can be used to investigate the effects of any chemical on the receptor characteristics e.g., electrophysiology (patch-clamp single-channel recording), binding assays (see section 1), etc.

Ligand-Gated G-protein linked Receptors

(e.g., mACh, 5HT, $GABA_B$, Glutamate, Dopamine, etc)

Many cell surface receptors are coupled to G-proteins (GTP-binding protein). G-protein-coupled receptors are thought to have seven membrane spanning domains, and have been divided into 2 subclasses: those in which the binding site is in the extracellular domain e.g. receptors for glycoprotein hormones, such as thyroid stimulating hormone (TSH) and follicle stimulating hormone (FSH), and those in which the ligand-binding site is likely to be in the plane of the 7 transmembrane domains e.g. rhodopsin and receptors for small neurotransmitters (nACh, 5HT, glutamate-NMDA, GABA, Glycine) and hormones. All transduce their signal by conformational change activation of an associated G-protein (see section 17).

There are two main classes of G proteins, the heterotrimeric G proteins that associate with receptors of the seven transmembrane domain superfamily and are involved in signal transduction, and the small cytoplasmic G

proteins. The small G proteins are a diverse group of monomeric GTPases that include ras, rab, rac and rho and that play an important part in regulating many intracellular processes including cytoskeletal organisation and secretion. Their GTPase activity is regulated by activators (GAPs) and inhibitors (GIPs) that determine the duration of the active state. (see section 17), see for example Principles of Neural Science, (Kandel and Schwartz), Third Edition 1991.

18. G-proteins (GTP binding proteins)

GTP binding proteins are a superfamily of related proteins which bind to guanosine nucleotides (Kaziro Y., Itoh H., Kozasa T., Nakafuku M. and Satoh T. Ann. Rev. Biochem. (1991). They are found in an inactive form which is bound to GDP and an active form which is bound to GTP. Other proteins such as ligand bound receptors promote the exchange of GDP with GTP, activating the protein. G proteins are inactivated by hydrolysis of the GTP to GDP. This reaction is catalysed by the G protein itself but the rate of GTP hydrolysis can be influenced by interaction with other proteins. Activated G proteins regulates the activities of a large number of target proteins including adenylate cyclase, phospholipase C and ion-channels.

Heterotrimeric G proteins.

Heterotrimeric G proteins are a large family of GTPases which consist of an α, a β and a γ subunit. They are involved in signal transduction from receptor proteins in the plasma membrane to second messenger systems within the cell receptors that activate. Activation of a receptor (e.g. by ligand binding) activates the G protein by promoting the exchange of bound GDP with GTP. The presence of GTP in the active site causes the dissociation of the α subunit from the $\alpha\beta\gamma$ complex. The free α subunits is most active. Different α subunit subtypes interact with a wide variety of different target proteins including adenylate

cyclase, phospholipase C and ion-channels. The free βγ complex also has also been show to have some regulatory activity.

Small(p21) GTPases

These proteins consist of a single subunit similar to the α subunit of heterotrimeric G proteins. These include the RAS family of proteins the abnormal activity of which can contribute to tumour formation.

Other GTP binding proteins

Other members of the Guanosine nucleotide binding protein superfamily include GTP binding translation elongation factors and members of the Dynamin family of proteins.

Use of Recombinant G proteins

Expression of recombinant G proteins allows the biochemical properties of proteins identified by DNA sequencing to be studied and allows the isolation of large amounts of the proteins for structural and biochemical studies. It also allows the production of mutant proteins produced by site directed modification of cDNA sequences.

Active recombinant G proteins have been expressed in large amounts in bacterial and insect-cell/baculovirus systems. Expression of G proteins in cell free translation systems is a convenient way of producing small amounts of protein for biochemical studies. The addition of ^{35}S methionine to the *in-vitro* translation reaction results in the production of specifically labelled protein.

It is also possible to express the proteins in cultured cells and look for whole cell effects such as increased cell proliferation, increased DNA synthesis or changes in the activity of various enzymes.

The use of cell lines lacking G protein subunits.

Several cell lines have been isolated, or made using gene-disruption techniques, which lack particular G protein subunits. The most widely used of these is the cyc^- variant of the S49 mouse lymphoma cell line lack the $G_s\alpha$ subunit. It is possible to add back recombinant or purified G proteins to investigate their function. Purified or *in-vitro* translated protein can be added back to membrane preparations from the cell lines or the cells can be transfected with plasmid constructs which express the protein.

GTP binding assays

The nonhydrolyzable GTP analogue $^{35}S\gamma GTP$ will bind to most GTP binding proteins in the absence of any activator molecule. Purified or in-vitro translated G protein can be incubated with $^{35}S\gamma GTP$ and the reaction products passed through a nitrocellulose filter. Protein bound $^{35}S\gamma GTP$ will be retained on the filter and the activity measured (Carty D.J. and Iyengar R. (1994). Methods in Enzymology. **237**: 38-45.).

GTPγS activation

Conformational changes in G proteins and changes in subunit interaction can be studied by incubating the G protein with GTPγS which binds to, and irreversibly activates, the protein. Conformational changes in subunits and changes in subunit interaction alter the sites available for degradation by trypsin. The tryptic fragments of radio-labelled protein can be run on a SDS PAGE gel and visualised by autoradiography. Subunit interaction can also be studied by looking as sedimentation rates during ultra centrifugation and by using chemical crosslinking agents (Audigier Y. (1994). Methods in Enzymology. **237**: 239-254.

Activation of other proteins as a result of G protein activation

G proteins in cell extracts can be activated by incubation with GTPγS and the activities of possible downstream target proteins such as adenylate cyclase and phospholipaseC measured.

Receptor stimulated GTP binding and GTP hydrolysis.

Receptor stimulated binding of the radio-labelled non-hydrolyzable GTP analog ^{35}SγGTP can be used to show if the addition of a receptor ligand leads to the activation of a G protein (Wieland T and Jakobs K.H (1994) Methods in Enzymology 237, 3 - 13). It is possible to study the activation of endogenous G proteins or to use a membrane preparation lacking particular G proteins and add back a purified or recombinant G protein.

^{35}SγGTP is added to a reaction mix containing a membrane preparation of the cells being studied. After incubation at 37°C for an appropriate length of time the reaction is stopped. The reaction mix is then passed through a filter which binds protein of membrane. The amount of radioactivity incorporated into the protein/membrane fraction is then measured. The amount of radioactivity incorporated in the presence and absence of candidate receptor ligand molecules can then be compared.

As an alternative to measuring the binding of ^{35}SγGTP it is possible to measure GTPase activity. Activation of a G protein by a ligand bound receptor results in an increase in GTP hydrolysis activity. This is more often a result of increasing the rate of exchange of GDP with GTP rather than an increasing the rate of hydrolysis of bound GTP. γ^{32}P GTP is added to a reaction mix containing a membrane preparation of cells and the amount of ^{32}P released from the labelled GTP is measured.

See also Methods in Enzymology, 195:pp171-474.

SEQ ID NO.	IDNumber	Class	Chr	Feature	AccNo	Name of match
1	NPS1	GNL	2	4-133	X54648	Frizzeled gene
2	NPS2	GNL	2	68-345	D17389	Ryanodyne receptor
3	NPS3	GNL	2	1-354	M23412	Muscarinic acetylcholine receptor. Genomic AC006938 intron.
4	NPS4	GNL	2	9-277	U91980	Tpr homolgue
5	NPS5	GNL	2	1-587	X61209	Type II topoisomerase
6	NPS6	GNL	2	270-408	U22439	Neuron surface antigen 2
7	NPS7	GNL	2	30-461INV	Y13272	Indora
8	NPS8	GNL	2	1-267	L03209	GDP dissociation inhibitor homologue
9	NPS9	GNL	2	345-583INV	L17340	germline transcription factor gene
10	NPS10	GNL	2	1-480	AB003784	Histone H3
11	NPS11	GNL	2	179-360	AA699128	EST matching 5' of V-ATPase C subunit
12	NPS13	GNL	2	226-409	U94702	MtPolB
13	NPS14	GNL	2	110-191	L13305	AND 396-472,integrin beta subunit (beta neu)
14	NPS15	GNL	2	1-432	X57484	tra-2 gene
15	NPS16	GNL	2	61-276	X15805	EF2 Translation factor
16	NPS18	GNL	2	1-532	X15008	49bp upstream of TU-36B gene, cytochrome b related protein.
17	NPS19	GNL	2	250-536	M29602	G0 protein alpha subunit homolog class II
18	NPS20	GNL	2	81-476	L13255	Lachesin
19	NPS21	GNL	2	119-457inv	U63556	larval serum protein 1 beta subunit
20	NPS22	GNL	2	121-417	AF027300	418-481intron,481-577exon.Positive transcription elongation factor b
21	NPS23	GNL	2	1-577	X84681	organellar-type Ca-ATPase gene.
22	NPS26	GNL	2	1-534	X71866	GTP-binding protein.
23	NPS27	GNL	2	1-523	M23094	Intron of G protein alpha subunit gene
24	NPS28	GNL	2	19-215	AF041048	AA246996 est match.EST matches CD39-like NTPase gene
25	NPS30	GNL	2	387-473inv	AF071417	phosphatidylinositol 4-phosphate 5-kinase, sktl1
26	NPS31	GNL	2	1-319	S55886	rbp9
27	NPS32	GNL	2	1-493	L34276	manganese superoxide dismutase (mnSOD)
28	NPS33	GNL	2	233-377inv	D84313	rab2
29	NPS34	GNL	2	1-63inv	AF003826	myosin V
30	NPS35	GNL	2	325-528	U09369	ribonucleoside-diphosphate reductase large subunit gene
31	NPS36	GNL	3	234-271	U95821	transmembrane GTPase (fzo)
32	NPS38	GNL	3	322-450	U00669	mitochondrial single-stranded DNA-binding protein
33	NPS39	GNL	3	14-385	X52846	RM62
34	NPS40	GNL	3	1-422inv	AF069297	pterin-4a-carbinolamine dehydratase gene
35	NPS41	GNL	3	329-346	Y09065	330-414 intron, 415-511exon.cytochrome c oxidase subunit Va preprotein
36	NPS42	GNL	3	1-283	M22428	Ubiquitin

38	NPS44	GNL	3	1-380	M17719	Intron of Rhodopsin 4 and M17730
39	NPS45	GNL	3	1-449	U27561	TipE
40	NPS46	GNL	3	463-528	X99665	mitochondrial ATPase coupling factor 6. Match on EST AI405330
41	NPS47	GNL	3	1-246	K01294	heat shock locus 87C1: proximal gene, 3' end.
42	NPS48	GNL	3	221-318	U73160	AA440389 EST matching Dros fas gene
43	NPS49	GNL	3	15-95	M32141	AI297861 1st EST in 8 contig matches 49-kilodalton phosphoprotein gene
44	NPS50	GNL	3	231-293	M21159	Tcp-1
45	NPS51	GNL	3	1-349	V00213	Hsp70
46	NPS52	GNL	3	1-241	U59923	glutamyl-prolyl-tRNA synthetase gene,
47	NPS53	GNL	3	225-237	D16257	238-333 intron, 334-499 exon ribosomal protein S4
48	NPS54	GNL	3	1-462	X73216	Rlb1
49	**NPS55**	**GNL**	**3**	**1-164inv**	**U62005**	**Rel/NF-kappa B homolog (Relish)**
50	NPS56	GNL	3	1-207inv	X07311	HSP2
51	NPS57	GNL	3	15-438	X54061	205K microtubule-associated protein (MAP)
52	NPS58	GNL	3	1-80inv	J01102	HSP68
53	**NPS59**	**GNL**	**3**	**1-450**	**Y10015**	**anon-66Da gene**
54	NPS60	GNL	3	56-187	M63792	RAD6
55	NPS61	GNL	3	391-465	U28966	Septin 2
56	NPS62	GNL	3	1-514	M98351	fructose 1,6 bisphosphate aldolase gene,
57	NPS63	GNL	3	46-251inv	U01035	Bottleneck gene
58	NPS64	GNL	3	49-450	U38238	HLH106
59	**NPS65**	**GNL**	**3**	**328-581**	**AB004232**	**DAD gene**
60	NPS66	GNL	3	1-436	U22176	15bp upstream of Brother gene on AC005557
61	NPS67	GNL	3	46-176	M90755	Transcriptional repressor protein Aef-1
62	NPS68	GNL	3	224-298	Y07908	Match to EST AI292767. This then matches serine/threonine protein kinase.
63	NPS69	GNL	3	1-531	M32311	Fascilin 1
64	NPS70	GNL	3	1-421inv	X03889	HSP23
65	NPS71	GNL	3	548-882inv	Y12861	bifunctional ATP sulfurylase/APS kinase.
66	NPS72	GNL	3	83-135	U12010	putative serine/threonine protein kinase (nemo)
67	NPS73	GNL	3	1-357	U20554	UDP-glucose:glycoprotein glucosyltransferase mRNA
68	NPS74	GNL	3	1-20bp	U87925	Cbl gene confirmed by match to EST AA441040
69	NPS75	GNL	3	468-539	U23485	Guanylate cyclase. Match found via EST AA392994
70	NPS76	GNL	3	1-547	Y11349	UbcD4
71	NPS77	GNL	3	1-163	U09374	SNAP
72	NPS78	GNL	3	1-104inv	U62388	chromatin assembly factor 1 p55 subunit
73	NPS79	GNL	3	374-518inv	AB007692	Elongin B
74	NPS80	GNL	3	1-231	L06861	232-401 intron, 402-590 exon matching TAF110

76	NPS82	EST	2	509-591	AA202837	hypothetical yeast/arabidopsis/prot and mouse EST
77	NPS83	EST	2	166-393	AI293734	
78	NPS84	EST	2	261-377inv	AA202757	Match to Human EST
79	NPS86	EST	2	1-247	AA696498	
80	NPS87	EST	2	100-646	AA950073	
81	NPS89	EST	2	1-50inv	AA695104	
82	NPS91	EST	2	1-427	AA942153	
83	NPS92	EST	2	42-334	AA540352	
84	NPS93	EST	2	115-162	AI238523	
85	NPS97	EST	2	1-69inv	AI260872	EST matches mouse signalling factor U29156
86	NPS98	EST	2	5-77bp	AA801728	
87	NPS99	GNL	2	228-675	AF053083	Drosophila SMT3 gene
88	NPS100	EST	2	1-210inv	AA439866	
89	NPS105	EST	2	31-590	AA820803	Poss related to human aldolase
90	NPS106	EST	2	30-478	AA803545	AA697132 match to frog/human MSS1
91	NPS108	EST	2	76-178	AA438591	
92	NPS109	EST	2	1-169	AA979551	also AA567400
93	NPS111	EST	2	138-414	AA439261	Match to Rat EST
94	NPS113	EST	2	7-354	AI107509	
95	NPS114	EST	2	1-48bp	AA540348	
96	NPS115	EST	2	1-311inv	AA735555	
97	**NPS118**	**EST**	**2**	**1-582**	**AI064020**	**match to AC005646. 26bp 5' to EST match area. AI542218/AI25740 765bp. Def SEC61 homologue**
98	**NPS119**	**EST**	**2**	**7-170**	**AA263700**	**also AA978721**
99	NPS120	EST	2	364-583	AA941785	also AA695548
100	**NPS121**	**EST**	**2**	**1-260 and 562-645**	**AA802928**	**also AA817115**
101	NPS122	EST	2	1-395	AA539001	
102	NPS123	EST	2	1-35inv	AA735863	Poss. related to human death assoc prot 3 X83544
103	NPS125	EST	2	68-195 and 475-621	AA941860	
104	NPS127	EST	2	1-210inv	AA246460	
105	NPS128	EST	2	66-593	AA141928	
106	NPS131	EST	2	1-332	AA979014	
107	NPS134	EST	2	52-475	AA817254	
108	NPS137	EST	2	1-37bp	AA536262	1209bp EST contig.AA948897. AA539274, AA392320. Poss glycogen synthase
109	**NPS139**	**GNL**	**2**	**35-86 and 475-581**	**AF113612**	**Drosophila Aspartate ligase**
110	NPS140	EST	2	368-636	AA390587	

112	NPS142	EST	2	31-460	AA941359	
113	NPS143	EST	2	65-299	AA201303	also AA541066
114	NPS144	EST	2	538-581inv	AA698119	Matc to Human gycerol-3-phosphate dehydrogenase
115	NPS145	EST	2	111-549	AA696174	
116	NPS146	EST	2	107-243	AI064230	Also AA263288. Match to Mouse proteasome subunit
117	NPS147	EST	2	1-212 and 276-382	AI106957	1756bp EST contig. Also AA391125, AA567307, AA735971
118	NPS149	EST	2	1-107inv	AI114218	Also AA820473. (AF034644) putative cytochrome bc-1 complex core protein [Haematobia irritans irritans]
119	NPS150	EST	2	19-102 and 115-485	AA978449	Also AA940834. 103-114 gap of unknown length
120	NPS152	EST	2	182-362	AA802905	
121	NPS154	EST	2	235-279 and 376-452	AI259166	Also AI296787. Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex precursor (human)P10515
122	NPS155	EST	2	1-238inv	AA951193	
123	NPS156	EST	2	326-482	AA696743	Also AA803977
124	NPS157	EST	2	11-512	AA990758	Also AA246427. 975bp contig
125	NPS158	EST	2	1-406	AA697797	
126	NPS159	EST	2	1-29inv	AA802206	1341contig. AA202662, AA801949, AA942041
127	NPS160	EST	2	344-592	AA978904	
128	NPS161	EST	2	101-223 and 292-551	AA202366	
129	NPS162	EST	2	103-468	AA950164	
130	NPS163	EST	2	23-98 and 102-602	AA952159	99-101 gap of unknown length.Match to mouse EST
131	NPS166	EST	2	102-512	AA392519	also AA695318 and AA441243.758bp contig.
132	NPS168	EST	2	304-541	AI515517	also AI404462. Poss Ras related protein
133	NPS169	EST	2	191-387	AA698481	
134	NPS170	EST	2	451-606inv	AA803082	2166bp EST contig. Poss. Alt splice. AA941565, AA820668, AA978815 and AA697381
135	NPS1067	EST	2	1-570	AI405762	Seq.sim to hypothetical prots from arabidopsis and C. elegans
136	NPS173	EST	2	1-38inv	AA391495	1135bp contig. AA439145 and AA949325. Match to mouse EST
137	NPS174	EST	2	353-476	AA942305	
138	NPS178	EST	2	72-391	AA951839	also AA979603
139	NPS179	EST	2	1-112	AI386817	also AI404737
140	NPS180	EST	2	435-475	AA438658	
141	NPS181	EST	2	31-212	AI106794	also AI107315
142	NPS1068	EST	2	1-228	AI403747	
143	NPS188	EST	2	1-272	AA802791	also AA390699
144	NPS189	EST	2	1-190	AA949990	also AA246423

146	NPS191	EST	2	202-472	AA978927	poss. PCF11p homolog (Homo sapiens)
147	NPS192	EST	2	84 -318	AA541084	also AA538937
148	NPS195	EST	2	390-509	AA951890	RIR2_mouse ribonucleoside- diphosphate reductase m2 chain
149	NPS198	EST	2	1-140	AA439230	
150	**NPS199**	**EST**	**2**	**3-522**	**AA948907**	**also AA942191**
151	NPS200	EST	2	25-76inv	AA802379	also AA246624
152	NPS1069	EST	2	9-100 AND 179-411	AI404485	
153	NPS1070	EST	2	60-449	AI108647	
154	NPS1071	EST	2	1-49inv	AA951902	other ESTs inc AA949796
155	**NPS204**	**GNL**	**2**	**1-489**	**AF143860**	**Drosophila RanGap gene**
156	**NPS205**	**EST**	**2**	**1-278**	**AA940865**	**Xenopus/ human chromosomal assembly protein(U1367)**
157	NPS206	EST	2	117-263	AA803314	also AA941391. Human B-cell receptor associated protein.
158	**NPS207**	**EST**	**2**	**209-405**	**AA201448**	**856bp contig.AA438721 and AA247046**
159	NPS209	EST	2	37-243	AA696343	also AA696180. Match to human/ C. elegans calponin
160	NPS210	EST	2	261-580inv	AA540783	
161	NPS211	EST	2	26-267 and 336-459	AA695850	also AA698310. FKB4_RABIT P59 PROTEIN
162	NPS212	EST	2	1-224	AA441346	also AA390646 and AA696470. 1677 contig
163	NPS213	EST	2	1-514	AI064375	
164	NPS216	EST	2	181-299	AA540197	also AA695503 and AA941503.732bp contig
165	NPS217	EST	2	167-212inv	AA979442	also AA392418
166	NPS218	EST	2	89-159	AA536378	also AA949458
167	NPS219	EST	2	1-570	AI515537	Genomic AC004345. Also AI062109. 50bp upstream of EST.
168	NPS220	EST	2	1-184	AA390646	1705bp contig with AA440523 and AA696470
169	NPS225	EST	2	1-104 and 310-467	AI064169	also AA816652
170	NPS226	EST	2	1-288	AA439345	802contig with AA949877 and AA439626
171	NPS227	EST	2	1-350	AA979503	181bp upstream of EST Genomic AC005452
172	NPS228	EST	2	1-93 and 170-446	AI293141	
173	NPS229	EST	2	12-244	AI107445	also AA390813
174	NPS233	EST	2	12-478	AA390942	
175	NPS235	EST	2	11-103 and 296-389	AA802688	Poss 10k HSP
176	NPS236	EST	2	1-414	AA392415	
177	NPS239	EST	2	1-22bp	AA695619	
178	NPS240	EST	2	399-542	AA142132	
179	NPS241	EST	2	366-520	AA536537	

181	NPS243	EST	2	186-593	AA441247	also AA820771
182	**NPS244**	**EST**	**2**	**318-431**	**AA202196**	**1942bp contig with AI108811, AA950029, AA202725, AA440491, and AA697007. Match to mouse EST**
183	NPS245	EST	2	83-319inv	AI064123	also AA263284. Match to human androgen induced prostate proliferative shutoff assoc. protein.
184	NPS247	EST	2	1-89bp	AA441173	
185	NPS250	EST	2	65-414inv	AA440852	also AA541034
186	NPS251	EST	2	2-131	AI062640	
187	NPS252	EST	2	1-77inv	AA695507	Poss. match to Rat cytochrome C
188	NPS254	EST	2	89-251	AA736186	also AA801973. Poss. match to horse Thioredoxin
189	NPS255	EST	2	1-417	AA697603	also AA801716
190	NPS256	EST	2	1-528	AA950741	
191	NPS257	EST	2	1-53bp	AI063204	887bp contig with AA697347 and AA201878
192	NPS258	EST	2	1-44bp	AA441029	
193	NPS259	EST	2	1-157	AI114266	1141bp contig with AA949325, AA735675 and AA391495. Poss.match to human GMP synthase
194	NPS260	EST	2	1-562	AA951648	1340bp contig with AA539581, AA802940 and AA263326
195	NPS261	EST	2	26-137 and 360-422	AA391135	Match to SEC61, different area to NPS118
196	NPS262	EST	2	1-124	AA696531	C.elegans pro7, Z66519/ mouse EST
197	NPS265	EST	2	442-549	AI124332	
198	NPS266	EST	2	52-382	AA949873	
199	NPS1073	EST	2	1-167	AI133902	see also AC006562 poss phosphate transporter
200	NPS269	EST	2	1-550	AI403609	Genomic AC005129, 420bp upstream of EST
201	NPS271	EST	2	299-375	AA391470	
202	**NPS272**	**GNL**	**2**	**37-77bp**	**AF085601**	**Drosophila inorganic pyrophosphatase NURF-38**
203	NPS273	EST	2	1-76inv	AA696584	
204	**NPS275**	**EST**	**2**	**1-319**	**AA439099**	**1132bp contig with AA949325 and AA940848 poss. GMP synthase (human)**
205	NPS276	EST	2	21-377	AA695424	
206	NPS277	EST	2	152-590inv	AA440949	
207	NPS278	EST	2	132-312	AI062455	also AA440915
208	NPS279	EST	2	68-311	AA816432	
209	NPS281	EST	2	1-258	AA979191	Match to human CGI-28
210	**NPS283**	**EST**	**2**	**2-318**	**AA391495**	
211	NPS285	EST	2	1-89bp	AA441636	AA820540 and AA817484. Alt splice
212	NPS1075	EST	2	59-488	AI295363	
213	NPS288	EST	2	51-170	AI114059	also AA941565

215	NPS290	EST	2	378-471	AA950084	also AA978669
216	NPS291	EST	2	20-236 and 292-439	AI062945	
217	NPS293	EST	2	1-312	AA440345	2293bp EST contig. Poss. human cleavage and polyadenylation specificity factor, 160 kd subunit.AA201536,AA539993,AA942332, AA979174 ,AA202096
218	**NPS294**	**EST**	**2**	**10-501**	**AA696930**	
219	NPS295	EST	2	8-437inv	AA440135	
220	NPS296	EST	2	75-157	AI063979	also AA802032
221	NPS297	EST	2	1-144	AA699194	
222	NPS298	EST	3	507-547	AA441233	also AA392152
223	NPS299	EST	3	1-79inv	AA438352	33% over 113 AA Plant oxygenase
224	NPS300	EST	3	480-534	AI455428	
225	NPS301	EST	3	11-190inv	AA246916	Rat Mitochondrial import receptor
226	NPS302	EST	3	233-348	AA392258	
227	NPS304	EST	3	1-41inv	AI296848	Prob. 40-kDa V-ATPase subunit (mam)
228	NPS305	EST	3	255-354	AI388389	
229	NPS306	EST	3	335-448inv	AA441471	also AA540182. 52% over 107 AA like Bov/Hum/Mouse RHO GDP-dissoc. ihibitor 1
230	NPS307	EST	3	22-242	AA439855	also AA567284.
231	NPS308	EST	3	1-141 and 397-446inv	AA941606	
232	NPS310	EST	3	209-435	AA392324	
233	**NPS311**	**EST**	**3**	**1-393inv**	**AA264796**	
234	NPS312	EST	3	1-152	AA540030	Poss rat calcium binding prot.
235	NPS313	EST	3	85-596	AI109898	
236	NPS314	EST	3	365-473	AI259723	
237	NPS316	EST	3	1-141	AI294469	
238	NPS317	EST	3	145-325	AA140945	
239	NPS318	EST	3	1-331	AI259816	Related to Epsin (Hum)
240	**NPS322**	**EST**	**3**	**209-433**	**AA141103**	
241	NPS323	EST	3	1-98inv	AA246767	also AA141059
242	NPS324	EST	3	1180239inv	AA441468	also AA142226. 42% over 128 AA like C. elegans prot.Z66496
243	NPS327	EST	3	1-82inv	AA247070	1366bp contig with AA567381, AA568013, AA540724. C. elegans prot/human EST
244	NPS328	EST	3	433-469	AA802401	Prob. Alg2, glycosyltransferase hom./ Mouse MER 5
245	NPS330	EST	3	1-96inv	AI135263	Alt splice
246	NPS331	EST	3	243-489	AA695904	
247	NPS334	EST	3	1-317	AA246386	also AA541060
248	NPS335	EST	3	311-427	AA264961	57% over 82AA like mouse/ human Thioredoxin

250	NPS338	EST	3	74-276 and 344-438	AA263803	
251	NPS339	EST	3	3-166inv	AA202200	also AA202128
252	NPS340	EST	3	1-48 inv	AA439530	
253	NPS341	EST	3	28-207	AI109459	Poss GPI-anchored protein(human)
254	NPS342	EST	3	471-506inv	AI109779	
255	NPS343	EST	3	147-247	AA141054	
256	NPS1061	EST	3	65-118inv	AA141365	
257	NPS345	EST	3	144-549	AI063643	
258	NPS346	EST	3	1-148	AI107445	also AA390813
259	NPS347	EST	3	1-75bp	AI297362	
260	NPS348	EST	3	96-230inv	AA392916	
261	NPS349	EST	3	1-47 and 145-317ivn	AA201223	2631bp contig. AA538867/AA439491/AA390780/AA390983/AA201661/AA391700/AA202007. Human 88
262	NPS351	EST	3	537-687	AI454966	
263	NPS352	EST	3	10-441	AA202767	715bpcontig with AA201231 and AA392823. 31% over 129AA like Rat Nup84 and Human 88 KDa nucleopore complex
264	NPS353	EST	3	3-40inv	AA201212	
265	NPS354	EST	3	1-33inv	AI404994	And AI260898. Alt splice
266	NPS356	EST	3	1-292	AA539914	1042bp contig with AA201959
267	NPS357	EST	3	36-454	AA440953	
268	NPS359	EST	3	145-253	AA264591	
269	NPS360	EST	3	47-380	AA539491	
270	NPS361	EST	3	202-381inv	AI403737	
271	NPS362	EST	3	270-443inv	AA567141	
272	NPS363	EST	3	1-478	AI134670	
273	NPS364	EST	3	413-535inv	AA263763	
274	NPS365	EST	3	1-99bp	AA568011	
275	NPS367	EST	3	64-449	AI107456	
276	NPS370	GNL	3	212-414	AF074957	Drosophila Karyopherin alpha
277	NPS371	EST	3	1-146	AI295205	and AA141054. Alt splice
278	NPS372	EST	3	8-382	AA567704	
279	NPS373	EST	3	74-224 and 297-344	AA539252	
280	NPS374	EST	3	1-347inv	AI260759	
281	NPS375	EST	3	1-77inv	AI260646	
282	NPS377	EST	3	160-306	AA202424	and AA264609.

284	NPS380	EST	3	322-573	AA802438	1030bp contig with AI063681
285	NPS381	EST	3	34-470inv	AA438500	
286	**NPS382**	**EST**	**3**	**14-153 and 216-348 ,419-445**	**AI456286**	
287	NPS383	EST	3	41-56 and 223-353	AI062265	1475 contig,AA694862 and AI064128.UNC51 ser/thr kinase (C.elegans)
288	NPS384	EST	3	1-429inv	AA247020	
289	NPS385	EST	3	1-143	AA264635	
290	NPS387	EST	3	58-491	AA201749	877bp contig with AA803278/ human hypothetical gene
291	NPS388	EST	3	1-162	AA392551	
292	**NPS389**	**EST**	**3**	**1-379**	**AA438539**	**EST contig1200bp, ORF at 5' end. Matches Human ERF 1. AA201773, AA263752 and AA439563**
293	NPS390	EST	3	297-447	AA141715	
294	NPS392	EST	3	80-161	AA695862	
295	NPS393	EST	3	2-132	AA201517	
296	NPS394	EST	3	176-239	AA202297	Vertebrate vacuolar ATPase
297	NPS395	EST	3	1-33bpinv	AA567483	
298	NPS396	EST	3	1-209 and 271-468	AA817479	
299	NPS397	EST	3	17-139	AA441327	ATG orf hits Rat (and other) sodium dependant dicarboxylate transporter AB001321 58% over 74 AA
300	NPS398	EST	3	1-391	AA698011	
301	NPS399	EST	3	67-207	AA951986	nucleolar protein p40 [Homo sapiens]
302	**NPS400**	**EST**	**3**	**1-186**	**AI295731**	**2018contig with AI258429, AA696170, AI109519, AA391348. MouseAPG-1 , hsp/osmotic shock gene**
303	NPS402	EST	3	1-82bp	AA201430	
304	NPS403	GNL	3	15-54bp	AF132912	Drosophila ARP gene. Match to EST matching ARP
305	NPS404	EST	3	1-140	AA541045	May be distantly related to cystatin
306	NPS406	EST	3	392-501	AA390337	
307	NPS407	EST	3	1-202 and 273-440inv	AA141555	Matches mouse/human ESTs
308	NPS408	EST	3	158-252	AA263730	A little like yeast hypothetical protein YOL124c
309	NPS409	EST	3	324-370 and 448-545	AI259832	and AA990765.Human Ubiquitin conjugating enzyme 12
310	NPS410	EST	3	75-483	AI514268	
311	NPS411	EST	3	1-435	AI293256	20bp 5' to EST on AC006562. Part of ORF similar to molybdenum cofactor biosynthesis protein A[Homo sapiens]
312	NPS412	EST	3	1-71 and 148-435	AA201987	Poss Asparaginase
313	NPS413	EST	3	12-408	AA540020	

315	NPS416	EST	3	15-404	AA201957	Mouse ESTs
316	NPS417	EST	3	1-353	AA695344	
317	NPS418	EST	3	1-450	AA441018	
318	**NPS419**	**EST**	**3**	**301-334inv**	**AA202301**	
319	NPS420	EST	3	245-393	AA735819	
320	NPS421	EST	3	1-147	AA440886	also AA695395. matches UMP kinase from C.elegans and bacteria
321	NPS422	EST	3	76-217	AA803640	859bp with AA803683 and AA803676. Human Ribosomal L28 protein
322	NPS424	EST	3	472-786 and 842-1073	AI257267	chick glycine cleavage system h protein
323	NPS425	EST	3	1-75bp	AA539327	
324	NPS426	EST	3	419-468	AI530922	
325	NPS427	EST	3	92-265	AI402854	
326	NPS428	EST	3	1-222 and 291-354	AA441362	40% like human/mouse proteasome subunit HsN3
327	NPS429	EST	3	1-219	AA202487	A little like hypothetical yeast protein YEY6
328	NPS430	EST	3	328-455	AA263590	39% over 61AA like human hRNP F
329	NPS431	EST	3	50-113	AA201496	57% over 50 AA like human oxoglutarate dehydrogenase
330	NPS432	EST	3	281-510	AA391430	
331	NPS434	EST	3	1-50 inv	AI292722	also AI534704
332	NPS435	EST	3	1-65inv	AA439393	
333	NPS436	EST	3	299-512inv	AA820797	also AA438876
334	NPS437	EST	3	1-52inv	AA697891	Homologue of Bovine gamma COP
335	NPS438	EST	3	1-31inv	AA696845	
336	NPS439	EST	3	1-384	AI259031	
337	NPS440	EST	3	1-82bp	AA803464	may match human hypothetical protein KIAA0258
338	NPS441	EST	3	169-489	AA539974	
339	NPS442	EST	3	1-46 and 432-524	AA941993	
340	NPS443	EST	3	43-431	AA803074	
341	**NPS444**	**EST**	**3**	**1-197 and 268-534**	**AA695870**	**Dog/rat/ Yeast signal peptidase 18kd subunit**
342	NPS445	EST	3	775-911	AA433251	
343	**NPS446**	**EST**	**3**	**718-1007**	**AI297203**	**1122bp contig with AA438815 and AI455195**
344	**NPS448**	**EST**	**3**	**1-106inv**	**AA694869**	**1632bp with AA735812, AA568063 and AA695306. mammalian transketolase**
345	NPS449	EST	3	1-99inv	AA392932	
346	NPS451	EST	3	1-260 and 310-510	AA391707	
347	NPS452	EST	3	45-141 AND 445-582	AI294564	Match to mouse EST
348	NPS453	GNL	3	345-460	AF152928	Drosophila karyophylin alpha 3.
349	NPS454	EST	3	1-177	AA540743	1129bp with AI064582, AI519458 and AA568024

351	NPS457	EST	3	1-73bp	AI062939	
352	NPS458	EST	3	927-1070	AI109224	
353	NPS459	EST	3	43-146	AA696728	Poss. isopentyl pyrophosphate isomerase
354	NPS460	EST	3	59-533	AI518328	
355	NPS461	EST	3	43-432	AA263622	
356	NPS463	EST	3	1-164	AA539661	Matches Human Proton ATPase like protein
357	NPS464	EST	3	1-152 and 214-579	AI388964	
358	NPS465	EST	3	365-462	AA438987	also AA264877. FXR1 mental retardation gene, Human
359	NPS466	EST	3	6-257	AA392117	
360	NPS468	EST	3	45-120 and 545-591	AA821194	987bp with AA736168. human 40s ribosomal protein s29
361	NPS469	EST	3	12-469	AA539752	
362	NPS473	EST	3	1-382 and 456-484	AA803203	
363	NPS476	EST	3	105-154	AA802887	also AA820871. Hypothetical C. elegans prot B0336.11
364	NPS477	EST	3	22-177	AA817394	V. similar to Dead box family of DNA helicases (initiation factors)
365	NPS479	EST	3	1-77inv	AI064638	
366	NPS480	EST	3	1-37bp	AA736157	also AA140746.Match tomouseEST
367	NPS482	EST	3	7-369	AA820427	
368	NPS483	EST	3	1-533	AA391736	1692bp with AA202259 and AA820861.
369	NPS484	EST	3	158-470	AA667184	
370	NPS486	EST	3	1-122	AA735277	1176bp with AA697907
371	NPS487	GNL	3	514-616	AF129080_1	Drosophila COP9 complex homolog subunit 1-2 DCH1-2
372	NPS489	EST	3	140-189	AA202581	Match to human EST
373	NPS490	EST	3	41-377	AA390775	
374	NPS491	EST	3	169-488	AA539898	
375	NPS492	EST	3	1-127	AA390463	
376	NPS493	EST	3	1-321inv	AA568061	1356bp with AA264532 and AA441674
377	NPS495	EST	3	102-311	AA141908	794bp with AA802528
378	NPS496	EST	3	1-100inv	AA539224	
379	NPS497	EST	3	431-622inv	AA246367	Hum ribosome S6 PK
380	NPS499	EST	3	46-319	AA817295	
381	NPS501	EST	3	35-383	AA439743	
382	NPS503	EST	3	1-264	AA441568	
383	NPS504	EST	3	337-408 and 479-568	AA247082	
384	NPS505	EST	3	1-321inv	AA201685	868bp contig with AA540405
385	NPS506	EST	3	83-218	AA540693	1450bp with AA441321, AA440080 andAA392794. Match to mouse EST

387	NPS508	EST	3	2-339	AA439667	
388	NPS509	EST	3	12-394	AA539198	
389	NPS510	EST	3	83-227 and 498-518	AA696927	Match to human citrin
390	NPS512	EST	3	360-501	AA438961	Human KIAA0160 gene
391	**NPS513**	**EST**	**3**	**53-260inv**	**AA735138**	
392	NPS514	EST	3	1-237inv	AI064414	
393	NPS515	EST	3	1-67inv	AA540712	809bp with AA440879 and AA440431.human cyclin G assoc. Kinase.
394	NPS516	GNL	3	596-682inv	AF132145	Drosophila damage-specific DNA binding protein DDBa p127 subunit
395	NPS517	EST	3	1-513	AF007159	
396	NPS518	EST	3	13-235inv	AI511691	
397	NPS519	EST	3	1-61inv	AA264883	also AA392712. Alt splice.
398	NPS520	EST	3	30-451	AA438399	821bp contig with AA439438
399	NPS521	EST	3	600-627	AA440272	1324bp contig with AA438941
400	NPS526	EST	3	65-483	AA264865	
401	NPS527	EST	3	1-246	AA263693	
402	NPS528	EST	3	137-160 and 379-475	AA698620	
403	NPS529	EST	3	1-51inv	AA391350	
404	NPS530	EST	3	100-499	AA392183	DNA J homologue
405	NPS531	EST	3	149-348 and 412-457	AA696390	mouse/human yeast/ub fusion protein 1
406	NPS532	EST	3	1-267inv	AA802961	816bp with AA817584
407	NPS533	EST	3	1-259inv	AA699045	Poss. Slug cDNA25
408	NPS534	EST	3	1-99bp	AA952055	1205bp with AA202358, AA202625 and AA951416. Siah binding protein 1(human)
409	NPS535	EST	3	1-610inv	AA142266	
410	NPS536	EST	3	52-534	AA696974	Matches Human CGI-37 protein
411	NPS537	EST	3	503-626 and 734-1069	AI532170	1565bp with AI544333 and AI062662. Definite transcription factor, MTF-1
412	NPS538	EST	3	442-569	AA567128	Match to mouse EST
413	NPS540	EST	3	1-79bp	AA950480	
414	NPS541	EST	3	2-360	AA950161	1217bp with AA950864 and AA950181
415	NPS542	EST	3	30-194 and 291-366	AA539625	882bp with AA202440 and AA390927
416	NPS543	EST	3	1-255	AA951297	
417	NPS544	EST	3	20-95 and 479-654	AA948996	781bp with AA541068 and AA950730
418	**NPS545**	**EST**	**3**	**1-378**	**AA941568**	

420	NPS547	EST	[illegible]	[illegible]	[illegible]	[illegible]
421	NPS548	EST	3	1-279	AA697191	1002bp with AA392404 and AA438791
422	NPS549	EST	3	650-689	AI518422	
423	NPS1065	EST	3	27-689	AI535025	
424	NPS551	EST	3	8-563	AA950826	
425	NPS553	EST	3	48-150	AA949934	Mam. Casein kinase
426	NPS555	EST	3	411-582	AI109292	
427	NPS556	EST	3	1-279inv	AA202259	1693bp with AA391736 and AA820861
428	NPS557	EST	3	131-647	AA142065	poss. succinate semialdehyde dehydrogenase
429	NPS558	EST	3	76-559	AA536402	
430-783	NPS559-920					Unknown
784	NPS921	GENO.	2	1-537INV	AC006073	In intron of gene coding for 246AA protein at 1663-44895 . No database matches
785	NPS922	GENO.	2	1-720	AC004299	In space before Drosophila Homologue of Human C-TAK 1 ser/thr kinase
786	NPS924	GENO.	2	1-599	AC004115	In space before gene coding for 372AA protein.16297-18074. No database matches.
787	NPS925	GENO.	2	1-581	AC004716	In intron of gene coding for 355AA protien. 51777-83843bp. No database matches
788	NPS926	GENO.	2	1-628	AC005889	No good predicted exons in this area.
789	NPS927	GENO.	2	1-536	L49408	space before gene coding for 401AA protein. 75943-77148bp. Sequence similarity to Mammlian glia maturation factor.
790	NPS928	GENO.	2	1-86inv	AC004296	In space before gene coding for 878AA protein. 39765-47183bp. Sequence similarity to Mouse G-protein
791	NPS929	GENO.	2	1-573inv	AC004306	No good predicted exons in this area.
792	NPS1077	GENO.	2	1-648	AC006472	In intron of gene coding for 1876AA protein at 62506-79351. Seq similarity hypothetical proteins from human and yeast
793	NPS931	GENO.	2	1-463inv	AC006092	In space before gene represented by ESTs AA990657 and AI294791.
794	NPS932	GENO.	2	1-519	AC006073	No good predicted exons in this area.
795	NPS933	GENO.	2	1-704inv	AC007176	In gene coding for 566AA protein.85106-110350bp(complement). Possible transcription factor
796	NPS935	GENO.	2	1-307inv	AC004423	In space before gene coding for a 702AA prootein at 47585-59400bp. Sequence similarity to Xenopus DNA repair protein XPGC
797	NPS336	GENO.	2	1-412inv	AC005448	Part of gene coding for 422AA protein at 34104-35373(complement). Strong sequece similarity to Drosophila Septin 2
798	NPS937	GENO.	2	1-478	AC004313	Possibly in 3'UTR of gene coding for 355AA protein at 39720-40727.Weak sequence similarity to potassium channel gene
799	NPS938	GENO.	2	1-489	AC004641	In intron of gene coding for 402AA protein at 101233-153110bp. Strong sequence similarity to Xenopus FLAP endonuclease.
800	NPS1078	GENO.	2	1-558inv	AC004306	No good predicted exons in this area.

801	NPS940	GENO.	2	1-[illegible]	[illegible]	**similarity to Mouse serine C-palmitoyltransferase**
802	NPS941	GENO.	2	1-544	AC005334	ESTs matching at 74650 (AA803646, AI518976, AI108114) Sequence similarity to U5 snRNP
803	NPS942	GENO.	2	1-201	AC004154	In intron of gene coding for a 292AA protein at 12210-30840bp. Sequence similarity to human geranylgeranyltransferase
804	NPS943	GENO.	2	1-524	AC004766	In intron of gene coding for a 1442AA protein at 40620-93241(complement). Sequence similarity to hypothetical C.elegans gene ZK1128.2.
805	NPS944	GENO.	2	1-621	AC004361	No good predicted exons in this area.
806	NPS945	GENO.	2	1-569inv	AC007185	No good predicted exons in this area.
807	**NPS946**	**GENO.**	**2**	**1-462**	**L49408**	**In space before gene coding for 401AA protein at 75943-77148bp. Sequence similarity to Mammilan glia maturation factor.**
808	NPS947	GENO.	2	1-233inv	AC005750	in intron of gene coding for an 1813AA protein at 51190-77775 (complement). Sequence similarity to Rat CPG2 protein
809	NPS948	GENO.	2	1-525	AC005269	No good predicted exons in this area.
810	NPS949	GENO.	2	1-531	AC005554	In intron of gene coding for 2355AA gene at 16134-35638. Sequence similarity to Rat Fatty acid synthase.
811	**NPS951**	**GENO.**	**2**	**1-443inv**	**AC004758**	**No good predicted exons in this area.**
812	NPS952	GENO.	2	1-498inv	AC005894	In intron gene coding for 196AA protein at 57158-63906bp. Weak sequence similarity to rat metalloprotease.
813	NPS954	GENO.	2	1-320	AC004564	In intron of gene codning foe a 156AA protein at 20899-39698bp. Sequence similarity to Arabidopsis Immunophilin
814	NPS956	GENO.	2	1-429	AC005716	In intron of gene coding for 732 AA protein at 63422-80946bp. Sequence similarity to Rat follistatin
815	NPS958	GENO.	2	1-71bp	AC007180	In intron of gene coding for 217AA protein at 81722-84603 (complement).No database matches
816	NPS1079	GENO.	2	1-75bp	AC004758	In intron of gene coding for a 945AA protein at 87648-113518. Strong sequence similarity to Human retinoblastoma binding protein 2
817	**NPS962**	**GNL.**	**2**	**1-116inv**	**AC001661**	**In Intron of gene coding for a 286AA protein at 49143-60866bp (complement). Is Dros Wing blister gene**
818	NPS963	GENO.	2	1-512inv	AC004758	In intron of gene coding for a 945AA protein at 87648-113518. Strong sequence similarity to Human retinoblastoma binding protein 2
819	NPS964	GENO.	2	1-54bp	AC004334	In space before gene coding for 433AA proetein at 21602-22903. No database matches.
820	NPS966	GENO.	2	1-557inv	AC005149	In intron of gene coding for 424AA protein at 70149-97938. No database matches.
821	NPS968	GENO.	2	1-202	AC005333	No good predicted exons in this area.
822	NPS970	GENO.	2	1-534inv	AC005334	In intron of gene coding for 309AA protein at 64276-77888bp. No database matches
823	NPS971	GENO.	2	1-438	AC006421	No good predicted exons in this area.
824	NPS972	GENO.	2	1-524INV	AC005443	No good predicted exons in this area.

						mammalian Uridine phophorylase.
826	NPS974	GENO.	2	1-535	AC005889	No good predicted exons in this area.
827	NPS975	GENO.	2	1-47bp	AC005130	In Intron of gene coding for 826AA protein at 3650 -13339bp (complement) ESTs AA949050 and AA817663 come from this gene. Sequence similarity to Helix-loop-helix genes
828	**NPS976**	**GENO.**	**2**	**1-551**	**AC005554**	**In intron of gene coding for 2355AA gene at 16134-35638. Sequence simlarity to Rat Fatty acid synthase.**
829	NPS977	GNL	2	1-100, 146-499	AF097364	**Drosophila Drongo gene**
830	**NPS978**	**GENO.**	**2**	**1-580**	**AC005554**	**Part of gene coding for 2355AA gene at 16134-35638. Sequence simlarity to Rat Fatty acid synthase.**
831	NPS979	GENO.	2	1-256inv	AC004722	Space before gene coding for 834AA protein at 33470-40630. Sequence similarity to bromodomain containing proteins.
832	NPS980	GENO.	2	1-406	AC003054	In intron of gene coding for a 822AA protein at 9312-46969. No database matches.
833	NPS982	GENO.	2	1-460	AC004280	No good predicted exons in this area.
834	NPS983	GENO.	2	1-99bp	AC004722	In intron of gene coding for a 289AA protein at 43804-61450bp. No database matches
835	NPS985	GENO.	2	1-178inv	AC001661	In intron of gene coding for 300AA protein at 30647-46841. Weak sequence similarity to Mouse surfeit gene
836	NPS986	GENO.	2	1-602	AC005269	No good predicted exons in this area.
837	NPS987	GENO.	2	1-562	AC004362	No good predicted exons in this area.
838	NPS988	GENO.	2	1-521	AC004370	In intron of gene coding for 1277AA protein at 40819-69834 (complement). Sequemce similarity to human nuclear transport receptor.
839	**NPS989**	**GENO.**	**2**	**1-619**	**L49408**	**No good predicted exons in this area.**
840	NPS991	GENO.	2	1-535inv	AC005447	In intron of gene coding for a 802AA protein at 28001-49228bp. No database matches.
841	NPS992	GENO.	2	1-342	AC004120	No good predicted exons in this area.
842	NPS993	GENO.	2	1-512	AC005454	In space before gene coding for a 399AA protein at 66560-68732bp. Sequence similarity to mitochondrial carrier protein genes.
843	NPS994	GENO.	2	1-515inv	AC005130	No good predicted exons in this area.
844	**NPS995**	**GENO.**	**2**	**1-499bp**	**AC005439**	**No good predicted exons in this area.**
845	NPS997	GENO.	2	1-565inv	AC005127	No good predicted exons in this area.
846	NPS998	GENO.	2	1-568	AC005889	3bp overlap with gene coding for1365AA protein at 67351-74867bp. ESTs AI106939 and AI296430 come from this gene
847	NPS999	GENO.	2	1-503	AC005558	In intron of gene coding for 1277AA protein at 40819-69834bp (complement). Sequemce similarity to human nuclear transport receptor.
848	NPS1000	GENO.	2	1-620	AC004351	In gene coding for 676AA protein at 34511-37955bp. Sequence similarity to mouse LUN gene.
849	NPS1001	GENO.	2	1-519inv	AC004766	In Intron of gene coding for 1467AA protein at 867-18363bp. Sequence similarity to Drosophila Lipase 3.

850	NPS1002	GENO.	2	1-80inv	AC005247	Valyl tRNA synthetase.
851	NPS1003	GENO.	2	1-370inv	AC005129	In intron of gene coding for a 1208AA protein at 56208-83122bp (complement). No database matches.
852	**NPS1004**	**GENO.**	**2**	**1-748inv**	**AC005894**	**No good predicted exons in this area.**
853	NPS1005	GENO.	2	1-535inv	AC005447	Part of gene coding for 239AA protein at 507-13551bp (complement). Sequence similarity to C. elegans gene acc. no. AF002196
854	NPS1006	GENO.	2	1-581inv	AC005643	In intron of gene coding for 242AA protein at 33006-40459bp (complement, incomplete sequencce). No database matches.
855	**NPS1007**	**GENO.**	**2**	**1-342**	**AC005554**	**Part of gene coding for 2355AA gene at 16134-35638. Sequence similarity to Rat Fatty acid synthase.**
856	NPS1009	GENO.	2	1-77inv	AC004532	In intron of gene coding for 1145AA protein at 5993-19843. Sequence similarity to C.elegans AF067608.
857	NPS1010	GENO.	2	1-496inv	AC007186	Space before gene coding for 351AA protein at 110375-111625bp. Sequence similarity to Human YL gene.
858	NPS1011	GENO.	2	1-582	AC007176	In intron of gene coding for 566AA protein at 84105-109350 (complement). Sequence similarity to zinc finger transcription factors.
859	NPS1012	GENO.	2	1-483	AC004423	In intron of gene coding for 604AA protein at 3052-8810bp.Sequence similarity to C.elegans AL021481 gene.
860	NPS1013	GENO.	2	1-560	AC005811	No good predicted exons in this area.
861	NPS1016	GENO.	2	1-596	AC005653	In intron of gene coding for 528AA at 40963-70180bp. Sequence similarity to C.elegans U40420
862	NPS1017	GENO.	2	1-539	AC004516	Part of gene coding for 1730AA protein at 50171-62324bp. Sequence similarity to C.elegans UNC89
863	NPS1019	GENO.	2	1-505inv	AC005285	In intron of gene coding for1142AA protein at 116605-128877bp. Sequence similarity to Guanine nucleotide exchange genes.
864	NPS1021	GENO.	2	1-504inv	AC007137	No good predicted exons in this area.
865	NPS1022	GENO.	2	1-191inv	AC005643	No good predicted exons in this area.
866	NPS1023	GENO.	2	1-468	AC004642	In intron of gene coding for 1296AA protein at 1037-28442(complement). Sequence similarity to putative lysophosphatidic acid acyltransferase [Mus musculus]
867	NPS1024	GENO.	2	1-578inv	AC005749	In intron of gene coding for 1481AA protein at 159-11694bp. Sequence similarity to KIAA0596 protein [Homo sapiens]
868	NPS1025	GENO.	2	1-598	AC007185	No good predicted exons in this area.
869	NPS1027	GENO.	2	1-634	AC004340	In space before gene coding for 864AA protein at 20219-29453 (complement). Good sequence similarity to Human sec24 homologue
870	**NPS1028**	**GENO.**	**2**	**1-415**	**AC005456**	**Space before gene coding for 450AA protein at 45629-48055bp (complement). Sequence similarity to Human GMP synthase.**
871	NPS1029	GENO.	2	1-198	AC004375	No good predicted exons in this area.

						NAT1/death assoc. prot
873	NPS1031	GENO.	2	1-495inv	AC004154	Space before gene coding for 292AA protein at 12210-30840bp. Sequence similarity to human rab geranylgeranyl transferase
874	NPS1032	GENO.	2	1-116	AC004328	In intron of gene coding for 1063AA protein at 63470-78557(complement). Has been predicted from Dros genomic AL035311 and has similarity to mouse BOP1
875	NPS1033	GENO.	2	1-581inv	AC005112	In intron of gene coding for 407AA protein at 3680-49217(complement). No database matches
876	NPS1034	GENO.	2	1-506inv	AC004367	Space before gene coding for 387AA protein at 3496-35348bp. No database matches
877	NPS1036	GENO.	2	1-411	AC005472	1st exon of gene coding for 626AA protein at 20859-27089bp. Sequence similarity to human NAT1/death assoc. prot
878	NPS1080	GENO.	2	1-492	AC007121	**gene coding for 365AA protein at 22511-56594 (complement). No database matches**
879	NPS1038	GENO.	3	1-291	L49405	In intron of gene codning for 665AA protein at 632-41487 (complement). Sequence similarity to human hypothetical SBBI03 protein
880	NPS1039	GENO.	3	1-454	AC004658	In intron of gene coding for 394AA protein (complement). No database matches
881	NPS1062	GENO.	3	1-376	AC007757	Also matches Dros. EST AA951801. Poss. Transcription factor.
882	NPS1044	GENO.	3	1-597	AC006091	In intron of gene coding for 564AA protein at 74249-97818bp. Sequence similarity to YII3_Yeast hypothetical protein.
883	NPS1045	GENO.	3	1-498	AC005720	No good predicted exons in this area.
884	NPS1046	GENO.	3	1-375	AC005814	**Space before gene coding for 562AA protein at 60626-68675. Sequence similarity to Rat NAB1.**
885	NPS1049	GENO.	3	1-486	AC004713	In intron of gene coding for 575AA protein at 10435-29297. Weak sequence similarity to CD36 genes.
886	NPS1050	GENO.	3	1-544	AC005813	No good predicted exons in this area.
887	NPS1051	GENO.	3	1-549	AC006936	No good predicted exons in this area.
888	NPS1052	GENO.	3	1-306inv	AC005425	No good predicted exons in this area.
889	NPS1053	GENO.	3	1-579inv	AC005720	In intron of gene coding for 394AA protein at 61401-93968. No database matches
890	NPS1056	GENO.	3	1-191inv	AC004266	In intron of gene coding for 931AA at 52045-70222bp (complement). Sequencw similarity to C. elegans Zinc finger protein. Drosophila EST AI259457 comes from this gene
891	NPS1059	GENO.	3	1-264	AC006936	No good predicted exons in this area.
892	NPS1063	GNL	3	488-536	AI062190	EST comes from Dros ferrochelatase
893	NPS1064	GENO.	3		AF104256	Sequence similarity to Human transcriptional co-activator CRSP150
894	NPS1076	EST	2	101-597	AI388606	1249bp contig with AI258281 and AI258326
895	NPS1081	GENO.	2	1-491	L49408	No good predicted exons in this area.
896	NPS1082	GENO.	2	1-475inv	AC005714	space before 390AA orf at 162294-163466. Hom to hum. death assoc prot 3
897	NPS1083	GENO.	2	1-461	AC004375	No good predicted exons in this area.
898	NPS1084	GENO.	2	12-419	AF000177	translational sequence similarity to human CaSm protein

SEQ ID NO.	IDNumber	Class	Chr	Feature	AccNo	Name of match
1	NPS0001	GNL	2	4-133	X54648	Frizzeled gene
2	NPS0002	GNL	2	68-345	D17389	Ryanodyne receptor
3	NPS0003	GNL	2	1-354	M23412	Muscarinic acetylcholine receptor. Genomic AC006938 intron.
4	NPS0004	GNL	2	9-277	U91980	Tpr homolgue
5	NPS0005	GNL	2	1-587	X61209	Type II topoisomerase
6	NPS0006	GNL	2	270-408	U22439	Neuron surface antigen 2
7	NPS0007	GNL	2	30-461INV	Y13272	Indora
8	NPS0008	GNL	2	1-267	L03209	GDP dissociation inhibitor homologue
10	NPS0010	GNL	2	1-480	X14215	Histone H3
11	NPS0011	GNL	2	179-360	AA699128	EST matching 5' of V-ATPase C subunit
12	NPS0013	GNL	2	226-409	U94702	MtPolB
13	NPS0014	GNL	2	110-191	L13305	AND 396-472,integrin beta subunit (beta neu)
14	NPS0015	GNL	2	1-432	X57484	tra-2 gene
15	NPS0016	GNL	2	61-276	X15805	EF2 Translation factor
16	NPS0018	GNL	2	1-532	X15008	49bp upstream of TU-36B gene, cytochrome b related protein.
18	NPS0020	GNL	2	81-476	L13255	Lachesin
19	NPS0021	GNL	2	119-457inv	U63556	larval serum protein 1 beta subunit
20	NPS0022	GNL	2	121-417	AF027300	418-481intron,481-577exon.Positive transcription elongation factor b
21	NPS0023	GNL	2	1-577	X84681	organellar-type Ca-ATPase gene.
22	NPS0026	GNL	2	1-534	X71866	GTP-binding protein.
23	NPS0027	GNL	2	1-523	M23094	Intron of G protein alpha subunit gene
24	NPS0028	GNL	2	19-215	AF041048	AA246996 est match.EST matches CD39-like NTPase gene
26	NPS0031	GNL	2	1-319	S55886	rbp9
27	NPS0032	GNL	2	1-493	L34276	manganese superoxide dismutase (mnSOD)
28	NPS0033	GNL	2	233-377inv	D84313	rab2
29	NPS0034	GNL	2	1-63inv	AF003826	myosin V
31	NPS0036	GNL	3	234-271	U95821	transmembrane GTPase (fzo)
32	NPS0038	GNL	3	322-450	U00669	mitochondrial single-stranded DNA-binding protein
34	NPS0040	GNL	3	1-422inv	AF069297	pterin-4a-carbinolamine dehydratase gene
35	NPS0041	GNL	3	329-346	Y09065	330-414 intron, 415-511exon.cytochrome c oxidase subunit Va preprotein
37	NPS0043	GNL	3	1-213	M28870	heterogeneous nuclear ribonucleoprotein isoforms, exon 1A.
38	NPS0044	GNL	3	1-380	M17719	Intron of Rhodopsin 4 and M17730
39	NPS0045	GNL	3	1-449	U27561	TipE
40	NPS0046	GNL	3	463-528	X99665	mitochondrial ATPase coupling factor 6. Match on EST AI405330
41	NPS0047	GNL	3	1-246	K01294	heat shock locus 87C1: proximal gene, 3' end.

42	NPS0048	GNL	3	221-318	U73160	AA440389 EST matching Dros fas gene
43	NPS0049	GNL	3	15-95	M32141	AI297861 1st EST in 8 contig matches 49-kilodalton phosphoprotein gene
44	NPS0050	GNL	3	231-293	M21159	Tcp-1
45	NPS0051	GNL	3	1-349	V00213	Hsp70. Poss EMPTY?
46	NPS0052	GNL	3	1-241	U59923	glutamyl-prolyl-tRNA synthetase gene,
47	NPS0053	GNL	3	225-237	D16257	238-333 intron, 334-499 exon ribosomal protein S4
48	NPS0054	GNL	3	1-462	X73216	Rlb1
50	NPS0056	GNL	3	1-207inv	X07311	HSP2
51	NPS0057	GNL	3	15-438	X54061	205K microtubule-associated protein (MAP)
52	NPS0058	GNL	3	1-80inv	J01102	HSP68
54	NPS0060	GNL	3	56-187	M63792	RAD6
55	NPS0061	GNL	3	391-465	U28966	Septin 2
56	NPS0062	GNL	3	1-514	M98351	fructose 1,6 bisphosphate aldolase gene,
57	NPS0063	GNL	3	46-251inv	U01035	Bottleneck gene
58	NPS0064	GNL	3	49-450	U38238	HLH106
60	NPS0066	GNL	3	1-436	U22176	15bp upstream of Brother gene on AC005557
61	NPS0067	GNL	3	46-176	M90755	Transcriptional repressor protein Aef-1
62	NPS0068	GNL	3	224-298	Y07908	Match to EST AI292767. This then matches serine/threonine protein kinase.
63	NPS0069	GNL	3	1-531	M32311	Fascilin 1
64	NPS0070	GNL	3	1-421inv	X03889	HSP23
65	NPS0071	GNL	3	548-882inv	Y12861	bifunctional ATP sulfurylase/APS kinase.
66	NPS0072	GNL	3	83-135	U12010	putative serine/threonine protein kinase (nemo)
67	NPS0073	GNL	3	1-357	U20554	UDP-glucose:glycoprotein glucosyltransferase mRNA
68	NPS0074	GNL	3	1-20bp	U87925	Cbl gene confirmed by match to EST AA441040
69	NPS0075	GNL	3	468-539	U23485	Guanylate cyclase. Match found via EST AA392994
70	NPS0076	GNL	3	1-547	Y11349	UbcD4
71	NPS0077	GNL	3	1-163	U09374	SNAP
72	NPS0078	GNL	3	1-104inv	U62388	chromatin assembly factor 1 p55 subunit
73	NPS0079	GNL	3	374-518inv	AB007692	Elongin B
74	NPS0080	GNL	3	1-231	L06861	232-401 intron, 402-590 exon matching TAF110
75	NPS0081	EST	2	1-314inv	AI259618	From genomic data.40bp upstream Prob. cytochrome B5. AC005641
76	NPS0082	EST	2	509-591	AA202837	hypothetical yeast/arabidopsis/prot and mouse EST
77	NPS0083	EST	2	166-393	AI293734	824bp EST contig with AI293545.
78	NPS0084	GNL	2	42-137inv	AJ249466	Dros DXI6 gene
79	NPS0086	EST	2	1-247	AA696498	
80	NPS0087	EST	2	100-646	AA950073	882bp EST contig with AI542046 and AI533154.Sequence similarity to Human PDI related protein

81	NPS0089	EST	2	1-50inv	AA695104	689bp EST contig with AA802812 and AI294978.
82	NPS0091	EST	2	1-427	AA942153	868bp EST contig with AA803815 and AI519003.
83	NPS0092	EST	2	42-334	AA540352	Also AI108950. poss. dehydrogenase
84	NPS0093	EST	2	115-162	AI238523	
85	NPS0097	EST	2	1-69inv	AI260872	EST matches mouse signalling factor U29156
86	NPS0098	EST	2	5-77bp	AA801728	830bp contig with AA263864 and forms part of mRNA AF184228.
87	NPS0099	GNL	2	228-675	AF053083	Drosophila SMT3 gene
88	NPS0100	EST	2	1-210inv	AA439866	
89	NPS0105	EST	2	31-590	AA820803	Other ESTs. evidence of Alt splice. Poss. aldose reductase
90	NPS0106	GNL	2	30-478	AF145307	Matches EST matching Dros. 26S proteasome regulatory complex subunit p48B
91	NPS0108	EST	2	76-178	AA438591	1199bp contig with AI107768 and AA697679.
92	NPS0109	GNL	2	1-169	AF147700	952bp EST contig AA979551/AA567400. Matches Dros Nebula gene. (EP line AQ254719-lethal ?)
93	NPS0111	EST	2	138-414	AA439261	1254bp contig with AI519697 and AA951893(polyA). Poss. zinc finger prot.
94	NPS0113	EST	2	7-354	AI107509	
95	NPS0114	GNL	2	1-48bp	AJ242855	ESTs AA735265 and AA540348. Match to Dros. amphiphysin. EP line AQ073590-lethal ?
96	NPS0115	EST	2	1-311inv	AA735555	1422bp contig with AI386481 and AI296915.
99	NPS0120	EST	2	364-583	AA941785	1115bp EST contig with AA695548 and AI513251(polyA).
101	NPS0122	EST	2	1-395	AA539001	
102	NPS0123	EST	2	1-35inv	AA735863	Poss. related to human death assoc prot 3 X83544 (EP line AQ73716-lethal?)
103	NPS0125	EST	2	68-195 and 475-621	AA941860	
104	NPS0127	EST	2	1-210inv	AA246460	Other ESTs, evidence of Alt. splice
105	NPS0128	EST	2	66-593	AA141928	ORF from MGD3 retrotransposon X95908
106	NPS0131	EST	2	1-332	AA979014	
107	NPS0134	EST	2	52-475	AA817254	
108	NPS0137	EST	2	1-37bp	AA536262	2350bp EST contig. Poss glycogen synthase
110	NPS0140	EST	2	368-636	AA390587	988bp EST contig with AI543996
111	NPS0141	EST	2	55-110inv	AA979454	sim to human REC1L protein. Acc. X57303
112	NPS0142	EST	2	31-460	AA941359	
113	NPS0143	EST	2	65-299	AA201303	also AA541066 Other ESTs, evidence of Alt. splice.
114	NPS0144	EST	2	538-581inv	AA698119	799bp EST contig. Match to Human gycerol-3-phosphate dehydrogenase
115	NPS0145	EST	2	111-549	AA696174	
116	NPS0146	EST	2	107-243	AI064230	922bp contig with AA263288 andAI109084 . Match to fish proteasome subunit (EP line AQ73732-lethaL?)
117	NPS0147	EST	2	1-212 and 276-382	AI106957	1906bp EST contig. 328ATG ORF/PolyA. Poss ATPase
118	NPS0149	EST	2	1-107inv	AI114218	Also AA820473. (AF034644) putative cytochrome bc-1 complex core protein (Haematobia irritans irritans)

119	NPS0150	EST	2	19-102 and 115-485	AA978449	Also AA940834. 103-114 gap of unknown length. Matches Dros mRNA AF145655
120	NPS0152	EST	2	182-362	AA802905	EP line AQ74042-=lethal?
121	NPS0154	EST	2	35-279 and 376-45	AI259166	Also AI296787. Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex precursor (human)P10515
122	NPS0155	EST	2	1-238inv	AA951193	EP line AQ72909 lethal?
123	NPS0156	EST	2	326-482	AA696743	Also AA803977
124	NPS0157	EST	2	11-512	AA990758	Also AA246427. 975bp contig. Matches Dros mRNA AF145681 (EP line AQ73763-lethal?)
125	NPS0158	EST	2	1-406	AA697797	1113bp EST contig with AI260878 and AI946139.
126	NPS0159	EST	2	1-29inv	AA802206	1626bp EST contig.
127	NPS0160	GNL	2	1-222and 422-592	AF035251	Dros. Tripeptidyl peptidase. TppII (FBgn0020370)
128	NPS0161	EST	2	01-223 and 292-55	AA202366	Dros. Unknown gene AF132158 (EP line AQ254870-lethal?)
129	NPS0162	EST	2	103-468	AA950164	
130	NPS0163	EST	2	23-98 and 102-602	AA952159	99-101 gap of unknown length.Match to mouse EST
131	NPS0166	EST	2	102-512	AA392519	also AA695318 and AA441243.758bp contig.
132	NPS0168	EST	2	304-541	AI515517	also AI404462. Poss Ras related protein
133	NPS0169	EST	2	191-387	AA698481	961bp EST contig with AI402698. Sim. to other hypothetical proteins. EP line AQ73580-lethal?
134	NPS0170	GNL	2	451-606inv	AA803082	2097bp EST contig. Poss. Alt splice. EST match Dros Akap200 dif, insert point to NPS0288
135	NPS1067	EST	2	1-570	AI405762	Seq.sim to hypothetical prots from arabidopsis and C. elegans
137	NPS0174	EST	2	353-476	AA942305	EP line AQ73548-lethal?
138	NPS0178	EST	2	72-391	AA951839	979bp EST contig with AA979603 and AI519882.
139	NPS0179	EST	2	1-112	AI386817	also AI404737
140	NPS0180	EST	2	435-475	AA438658	951bp EST contig with AI546582
142	NPS1068	EST	2	1-228	AI403747	
143	NPS0188	EST	2	1-272	AA802791	also AA390699
145	NPS0190	EST	2	1-202	AA201161	Same EST contig as NPS180, diiferent insertion points
146	NPS0191	EST	2	202-472	AA978927	Also AA978887. poss. PCF11p homolog [Homo sapiens]
147	NPS0192	EST	2	84 -318	AA541084	930bp EST contig withAA538937. Poss. choline kinase
148	NPS0195	GNL	2	390-509	U09370	Matches EST matching Dros. ribonucleoside diphosphate reductase
149	NPS0198	EST	2	1-140	AA439230	
151	NPS0200	EST	2	25-76inv	AA802379	1135bp EST contig with AA246624 and AI294568.
152	NPS1069	EST	2	9-100 AND 179-411	AI404485	
153	NPS1070	EST	2	60-449	AI108647	
154	NPS1071	EST	2	1-49inv	AA951902	other ESTs inc AA949796. Sequence sim. to human ARP2/3 protein compex subunit 34
157	NPS0206	EST	2	117-263	AA803314	1505bp EST contig. Evidence of alt. splice. Human B-cell receptor associated protein.
159	NPS0209	EST	2	37-243	AA696343	also AA696180. Match to human/ C. elegans calponin
160	NPS0210	EST	2	261-580inv	AA540783	659bp EST contig with AI456854.
161	NPS0211	EST	2	26-267 and 336-459	AA695850	1305bp EST contig with AA698310 and AI516896. FKB4_RABIT P59 PROTEIN

162	NPS0212	GNL	2	1-224	AF149795	Dros Drep2-no flybase
163	NPS0213	EST	2	1-514	AI064375	
164	NPS0216	EST	2	181-299	AA540197	also AA695503 and AA941503.732bp contig. Poss sialomucin
165	NPS0217	EST	2	167-212inv	AA979442	732bp contig with AA392418 and AI543860.
166	NPS0218	EST	2	89-159	AA536378	2393bp EST contig. Poss arsenate resistance protein ARS2(human)
167	NPS0219	EST	2	1-570	AI515537	Genomic AC004345. Also AI062109. 50bp upstream of EST.(EP line AQ073614-lethal?)
168	NPS0220	GNL	2	1-184	AF149795	Dros Drep2-no flybase
169	NPS0225	EST	2	1-104 and 310-467	AI064169	also AA816652
170	NPS0226	EST	2	1-288	AA439345	802contig with AA949877 and AA439626
171	NPS0227	EST	2	1-350	AA979503	806bp EST contig with AI512942.181bp upstream of EST Genomic AC005452
172	NPS0228	EST	2	1-93 and 170-446	AI293141	
173	NPS0229	EST	2	12-244	AI107445	955bp EST contig with AA390813
174	NPS0233	EST	2	12-478	AA390942	
175	NPS0235	EST	2	11-103 and 296-389	AA802688	Also AI134220. Poss 10k HSP
176	NPS0236	EST	2	1-414	AA392415	
177	NPS0239	EST	2	1-22bp	AA695619	
178	NPS0240	EST	2	399-542	AA142132	Dros mRNA AB010264
179	NPS0241	EST	2	366-520	AA536537	EP line AQ073930-lethal?
180	NPS0242	EST	2	26-303	AA264253	Also AI106910. Poss. SNF7 homologue
181	NPS0243	EST	2	186-593	AA441247	also AA820771
183	NPS0245	EST	2	83-319inv	AI064123	also AA263284. Match to human androgen induced prostate proliferative shutoff assoc. protein.
184	NPS0247	EST	2	1-89bp	AA441173	
185	NPS0250	GNL	2	1-347inv	AF168467	Dros dual specificity kinase DYRK2(no flybase)
186	NPS0251	EST	2	2-131	AI062640	1190bp EST contig with AI063780. EP line AQ074008-lethal?
187	NPS0252	EST	2	1-77inv	AA695507	1392bp EST contig. Poss.Alt splice. Poss. RNA binding prot.
188	NPS0254	EST	2	89-251	AA736186	1270bp EST contig with AA801973 and AI294493 . Poss. match to horse Thioredoxin. EP line AQ073432-lethal?
189	NPS0255	EST	2	1-417	AA697603	also AA801716. EP line AQ073594
190	NPS0256	EST	2	1-528	AA950741	
191	NPS0257	EST	2	1-53bp	AI063204	1079bp EST contig.
192	NPS0258	EST	2	1-44bp	AA441029	
193	NPS0259	EST	2	1-157	AI114266	EP line AQ073622-lethal? 1141bp contig with AA949325, AA735675 and AA391495. Poss.match to human GMP synthase
194	NPS0260	EST	2	1-562	AA951648	1340bp contig with AA539581, AA802940 and AA263326
195	NPS0261	EST	2	26-137 and 360-422	AA391135	Match to SEC61, different area to NPS118. EP line AQ072909 lethal?
196	NPS0262	EST	2	1-124	AA696531	Sim. to 2-hydroxyphytanoyl-CoA lyase (Human)

197	NPS0265	EST	2	442-549	AI124332	
198	NPS0266	EST	2	52-382	AA949873	
199	NPS1073	EST	2	1-167	AI133902	see also AC006562 poss phosphate transporter
200	NPS0269	EST	2	1-550	AI403609	Genomic AC005129, 420bp upstream of EST
201	NPS0271	EST	2	299-375	AA391470	
203	NPS0273	EST	2	1-76inv	AA696584	
205	NPS0276	EST	2	21-377	AA695424	Also AI295527
206	NPS0277	EST	2	152-590inv	AA440949	
207	NPS0278	EST	2	132-312	AI062455	also AA440915
208	NPS0279	EST	2	68-311	AA816432	
209	NPS0281	EST	2	1-258	AA979191	Match to human CGI-28
211	NPS0285	EST	2	1-89bp	AA441636	AA820540 and AA817484. Alt splice
212	NPS1075	EST	2	59-488	AI295363	
213	NPS0288	GNL	2	64-170	AF132884	Dros Akap200 also ep line AQ074077 lethal ?(no flybase)
214	NPS0289	EST	2	3-355 and 443-479	AA801691	also AA441008
215	NPS0290	EST	2	378-471	AA950084	1473bp EST contig.
216	NPS0291	EST	2	20-236 and 292-439	AI062945	798bp EST contig with AI405532 and AA264809. Some similarity to Human p47 protein.
217	NPS0293	EST	2	1-312	AA440345	2293bp EST contig. Poss. human cleavage and polyadenylation specificity factor, 160 kd subunit.AA201536,AA539993,AA942332, AA979174 ,AA202096
219	NPS0295	EST	2	8-437inv	AA440135	EP line AQ073560-lethal?
220	NPS0296	EST	2	75-157	AI063979	also AA802032
221	NPS0297	EST	2	1-144	AA699194	
222	NPS0298	EST	3	507-547	AA441233	also AA979182
223	NPS0299	EST	3	1-79inv	AA438352	33% over 113 AA Plant oxygenase
224	NPS0300	EST	3	480-534	AI455428	
225	NPS0301	EST	3	11-190inv	AA246916	Also AI532444. Rat Mitochondrial import receptor
226	NPS0302	EST	3	233-348	AA392258	Also AI259457
227	NPS0304	EST	3	1-41inv	AI296848	Prob. 40-kDa V-ATPase subunit (mam)
228	NPS0305	EST	3	255-354	AI388389	
229	NPS0306	EST	3	335-448inv	AA441471	also AA540182. 52% over 107 AA like Bov/Hum/Mouse RHO GDP-dissoc. ihibitor 1
230	NPS0307	EST	3	22-242	AA439855	2093bp EST contig. Evidence of Alt splice. seq. sim. to Human HYDROXYMETHYLBILANE SYNTHASE
231	NPS0308	EST	3	-141 and 397-446in	AA941606	also AA978838. Seq. sim to serine proteinase [Anopheles gambiae]
232	NPS0310	EST	3	209-435	AA392324	EP line AQ073800 lethal?
234	NPS0312	EST	3	1-152	AA540030	Also AI517519. Poss rat calcium binding prot. EP line AQ073037 lethal?
235	NPS0313	EST	3	85-596	AI109898	
236	NPS0314	EST	3	365-473	AI259723	

237	NPS0316	EST	3	1-141	AI294469	
238	NPS0317	EST	3	145-325	AA140945	Dros Fex retrotransposon
239	NPS0318	GNL	3	1-331	AF160975	Dros. Liquid facets gene. Epsin homologue. No flybase.
241	NPS0323	EST	3	1-98inv	AA246767	also AA141059
242	NPS0324	EST	3	1180239inv	AA441468	1167bp EST contig.
243	NPS0327	EST	3	1-82inv	AA247070	2486bp EST contig. Poss. translation factor.
244	NPS0328	EST	3	433-469	AA802401	Also AA802218. Prob. Alg2, glycosyltransferase hom./ Mouse MER 5
245	NPS0330	EST	3	1-96inv	AI135263	1514bp EST contig or 1119AS contig, ATG ORFS, no good hits.
246	NPS0331	GNL	3	243-489	Y14998	Dros Bip1 gene (no flybase)
247	NPS0334	EST	3	1-317	AA246386	1392bp EST contig with AA541060 and AI294651
248	NPS0335	EST	3	311-427	AA264961	1083bp EST contig with AI402225. 57% over 82AA like mouse/ human Thioredoxin
250	NPS0338	EST	3	74-276 and 344-438	AA263803	
251	NPS0339	EST	3	3-166inv	AA202200	1329bp EST contig. EP line AQ073952 lethal?
252	NPS0340	EST	3	1-48 inv	AA439530	Also AA951823
253	NPS0341	EST	3	28-207	AI109459	764bp EST contig with AA695738. Poss GPI-anchored protein(human)
254	NPS0342	EST	3	471-506inv	AI109779	
255	NPS0343	EST	3	147-247	AA141054	
256	NPS1061	EST	3	65-118inv	AA141365	
257	NPS0345	EST	3	144-549	AI063643	
258	NPS0346	EST	3	1-148	AI107445	1662bp EST contig. Sequence sim. to yeast proteins.
259	NPS0347	EST	3	1-75bp	AI297362	
260	NPS0348	EST	3	96-230inv	AA392916	EP line AQ025069-lethal? Evidence of Alt. splice
261	NPS0349	EST	3	1-47 and 145-317iv	AA201223	4020bp EST contig (11ESTs). Probable translation initiation factor
262	NPS0351	EST	3	537-687	AI454966	
263	NPS0352	EST	3	10-441	AA202767	901bpcontig with AI387427,AA201231 and AA392823. 31% over 129AA like Rat Nup84 and Human 88 KDa nucleopore complex
264	NPS0353	EST	3	3-40inv	AA201212	Also AI237918
265	NPS0354	EST	3	1-33inv	AI404994	And AI260898. Alt splice
266	NPS0356	EST	3	1-292	AA539914	1500bp EST contig. Some sim. to human nudix (nucleoside diphosphate linked moiety X)
267	NPS0357	EST	3	36-454	AA440953	705bp EST contig with AI518505
268	NPS0359	EST	3	145-253	AA264591	EP line AQ073779 -lethal?
270	NPS0361	EST	3	202-381inv	AI403737	
271	NPS0362	EST	3	270-443inv	AA567141	2042bp EST contig. Some seq. sim. to C.elegans hypothetical genes
272	NPS0363	EST	3	1-478	AI134670	Also AI108248(Alt. splice)
273	NPS0364	EST	3	413-535inv	AA263763	Also AA696172, some seq.sim. to mouse EX070
274	NPS0365	GNL	3	1-99bp	AF190745	EST AA568011 matches Dros. PolyU binding splicing factor.No flybase.
275	NPS0367	GNL	3	64-449	AF145627	Dros Scratch gene scrt (no flybase)

276	NPS0370	GNL	3	212-414	AF074957	Drosophila Karyopherin alpha
277	NPS0371	EST	3	1-146	AI295205	and AA141054. Alt splice
278	NPS0372	EST	3	8-382	AA567704	1466bp EST contig.EP line AQ025097 lethal7Seq. sim to mouse malate oxidoreductase.
280	NPS0374	EST	3	1-347inv	AI260759	
281	NPS0375	GNL	3	1-199inv	AF170829	Dros homolgue of mouse NP15.6 (neuronal)
282	NPS0377	EST	3	160-306	AA202424	1650bp EST contig.Matches full length mRNA AF184229. EP line AQ073843 lethal?
283	NPS0379	EST	3	300-379	AA802555	Also AI064221, evidence of Alt. splice.
284	NPS0380	EST	3	322-573	AA802438	1030bp contig with AI063681
285	NPS0381	EST	3	34-470inv	AA438500	Also AA735111. EP line AQ025197-lethal?
287	NPS0383	EST	3	41-56 and 223-353	AI062265	1475 contig.AA694862 and AI064128.UNC51 ser/thr kinase (C.elegans)
288	NPS0384	EST	3	1-429inv	AA247020	
289	NPS0385	EST	3	1-143	AA264635	Also AA816470.
290	NPS0387	EST	3	58-491	AA201749	877bp contig with AA803278/ human hypothetical gene
291	NPS0388	EST	3	1-162	AA392551	Also AA951287. EP line AQ072985
293	NPS0390	EST	3	297-447	AA141715	
294	NPS0392	EST	3	80-161	AA695862	
295	NPS0393	EST	3	2-132	AA201517	1104bp EST contig. Seq. sim. to Dros minidiscs gene
296	NPS0394	EST	3	176-239	AA202297	788bp contig matching complete mRNA AF132183. Seq. sim. toVertebrate vacuolar ATPase
297	NPS0395	GNL	3	1-33bpinv	AA567483	EST matches Dros protein tyrosine phosphatase L11253.
298	NPS0396	EST	3	1-209 and 271-468	AA817479	
299	NPS0397	EST	3	17-139	AA441327	ATG orf hits Rat (and other) sodium dependant dicarboxylate transporter AB001321 58% over 74 AA
300	NPS0398	GNL	3	1-391	AA698011	EST matches Dros Melted gene, AF205831
301	NPS0399	EST	3	67-207	AA951986	nucleolar protein p40 [Homo sapiens]
303	NPS0402	EST	3	1-82bp	AA201430	848bp EST contig with AI256930. seq. sim. to yeast mitosis protein.
304	NPS0403	GNL	3	15-54bp	AF132912	Drosophila ARP gene. Match to EST matching ARP
305	NPS0404	EST	3	1-140	AA541045	1850bp EST contig.May be distantly related to cystatin
306	NPS0406	EST	3	392-501	AA390337	Also AA990810. EP line AQ254714 lethal?
307	NPS0407	EST	3	-202 and 273-440in	AA141555	Matches mouse/human ESTs
308	NPS0408	EST	3	158-252	AA263730	Also AI531907(Alt splice). A little like yeast hypothetical protein YOL124c
309	NPS0409	EST	3	24-370 and 448-54	AI259832	916bp EST contig with AI946646 and AA990765. Good Seq. sim. toHuman Ubiquitin conjugating enzyme 12
310	NPS0410	EST	3	75-483	AI514268	
311	NPS0411	EST	3	1-435	AI293256	20bp 5' to EST on AC006562. Part of ORF similar to molybdenum cofactor biosynthesis protein A[Homo sapiens]
312	NPS0412	EST	3	1-71 and 148-435	AA201987	Also AI293141. Poss Asparaginase
313	NPS0413	GNL	3	14-408	AF104357	Dros DRONC Nedd2 like caspase

314	NPS415	EST	3	1-38bp	AA201670	
315	NPS416	EST	3	15-404	AA201957	Mouse ESTs
316	NPS0417	EST	3	1-353	AA695344	
317	NPS0418	EST	3	1-450	AA441018	Also AA264070. Matches lethal P-element insertion.AF1746829 7/12/99 poss. in dalmation, (no sequence)
319	NPS0420	EST	3	245-393	AA735819	
320	NPS0421	GNL	3	1-147	AA440886	EST matches Dros Dak1 AB025924 (no flybase)
321	NPS0422	EST	3	76-217	AA803640	672bp with AA803683 and AA803676. Human Ribosomal L28 protein
322	NPS0424	GNL	3	72-786 and 842-107	AI257267	722bp EST contig with AI947051.Match to Dros pumpless gene AF203725. Lethal known Dec1999
323	NPS0425	EST	3	1-75bp	AA539327	653bp EST contig with AI294865.
324	NPS0426	EST	3	419-468	AI530922	
325	NPS0427	EST	3	92-265	AI402854	
326	NPS0428	EST	3	1-222 and 291-354	AA441362	1218bp EST contig.40% like human/mouse proteasome subunit HsN3
327	NPS0429	EST	3	1-219	AA202487	1122bp EST contig with AA696163 and AA803708. A little like hypothetical yeast protein
328	NPS0430	EST	3	328-455	AA263590	1222bp EST contig. Seq. sim. to human hRNP F
329	NPS0431	EST	3	50-113	AA201496	2556bp EST contig. Seq sim to human oxoglutarate dehydrogenase
330	NPS0432	EST	3	281-510	AA391430	787bp EST contig with AA802523
331	NPS0434	EST	3	1-50 inv	AI292722	
332	NPS0435	GNL	3	1-65inv	AA439393	1828bp EST contig. Evidence of alt. splice. Part of Dros nemo gene. Different insert point to nps0072.
333	NPS0436	EST	3	299-512inv	AA820797	1851bp EST contig. Seq. sim. to human SEC63
334	NPS0437	GNL	3	1-52inv	AF191563	Dros coatomer protein. No flybase.
335	NPS0438	GNL	3	1-31inv	AA696845	875bp EST contig with AI456186.Matches Dros Klar. gene AF157066.EP line AQ254737 - lethal?
336	NPS0439	EST	3	1-384	AI259031	
337	NPS0440	EST	3	1-82bp	AA803464	Also AI064211(alt. splice).Seq. sim. to human hypothetical protein KIAA0258
338	NPS0441	EST	3	169-489	AA539974	842bp EST contig (evidence of alt. splice). EP line AQ025134-lethal?
339	NPS0442	EST	3	1-46 and 432-524	AA941993	
340	NPS0443	EST	3	43-431	AA803074	Also AI517810. Match to Dros mRNA(complete) AF145665
342	NPS0445	EST	3	775-911	AA433251	
345	NPS0449	EST	3	1-99inv	AA392932	819bp EST contig with AA392017.
346	NPS0451	EST	3	1-260 and 310-510	AA391707	1165bp EST contig with AI546088. Some similarity to human MSP2 protease.
347	NPS0452	EST	3	5-141 AND 445-58	AI294564	Also AI511708. Match to mouse EST
348	NPS0453	GNL	3	345-460	AF152928	Drosophila karyophylin alpha 3.
349	NPS0454	EST	3	1-177	AA540743	1129bp with AI064582, AI519458 and AA568024
350	NPS0456	EST	3	1-325	AA539054	Match to rat EST

351	NPS0457	EST	3	1-73bp	AI062939	Some sequence similarity to human ataxia 2
352	NPS0458	EST	3	927-1070	AI109224	
353	NPS0459	EST	3	43-146	AA696728	Poss. isopentyl pyrophosphate isomerase
354	NPS0460	EST	3	59-533	AI518328	
355	NPS0461	EST	3	43-432	AA263622	Also AI135048.Match to complete mRNA AF132157. EP line AQ254717-lethal?
356	NPS0463	EST	3	1-164	AA539661	1304bp EST contig. Matches Human Proton ATPase like protein
357	NPS0464	EST	3	1-152 and 214-579	AI388964	
358	NPS0465	GNL	3	365-462	AF205597	Dros Fragile X related gene. no flybase.
359	NPS0466	EST	3	6-257	AA392117	Also AA392507
360	NPS0468	EST	3	45-120 and 545-591	AA821194	1599bp EST contig . human 40s ribosomal protein s29
361	NPS0469	EST	3	12-469	AA539752	813bp EST contig with AI386513(alt splice).
362	NPS0473	EST	3	1-382 and 456-484	AA803203	1218bp EST contig.
363	NPS0476	EST	3	105-154	AI297317	Evidence of Alt splice
364	NPS0477	EST	3	22-177	AA817394	Also AA263804. V. similar to Dead box familly of DNA helicases (initiation factors)
365	NPS0479	EST	3	1-77inv	AI064638	
366	NPS0480	EST	3	1-37bp	AA736157	1462bp EST contig with AI109785 and AI114024. Seq. sim to 54TMp [Homo sapiens]
367	NPS0482	EST	3	7-369	AA820427	Seq sim. to 3-oxoacyl carrier protein synthase II [Arabidopsis thaliana]
368	NPS0483	EST	3	1-533	AA391736	2188bp EST contig.
369	NPS0484	EST	3	158-470	AA567184	1425bp EST contig.Evidence of Alt. splice. Good seq. sim. to human spliceosomal protein SAP 130
370	NPS0486	EST	3	1-122	AA735277	Same ESTs as NPS0361, different point of insertion
371	NPS0487	GNL	3	514-616	AF129080_1	Drosophila COP9 complex homolog subunit 1-2 DCH1-2
372	NPS0489	EST	3	140-189	AA202581	1210bp EST contig. Match to AMMECR1 protein [Homo sapiens]
373	NPS0490	EST	3	41-377	AA390775	Also AI292888
374	NPS0491	EST	3	169-488	AA539898	
375	NPS0492	EST	3	1-127	AA390453	Also AA951634
376	NPS0493	EST	3	1-321inv	AA568061	1356bp with AA264532 and AA441674
379	NPS0497	GNL	3	431-837inv	AF142061	Dros prot. kinase JIL-1
380	NPS0499	EST	3	46-319	AA817295	Seq. similarity to bestrophin [Homo sapiens].EP line AQ073104-lethal?
381	NPS0501	EST	3	35-383	AA439743	Also AI403533. Evidence of alt splice.
382	NPS0503	EST	3	1-264	AA441568	1164bp EST contig. EP line AQ025280-lethal?
383	NPS0504	EST	3	37-408 and 479-56	AA247082	
384	NPS0505	EST	3	1-321inv	AA201685	868bp contig with AA540405 and AI513192
386	NPS0507	EST	3	452-1044	AI295950	
387	NPS0508	EST	3	2-339	AA439667	also AI54395. Sequence similarity to human spermidine aminopropyltransferase.
388	NPS0509	EST	3	12-394	AA539198	
389	NPS0510	GNL	3	83-227 and 498-518	AA696927	1141bp contig with AI107574 and AA801718.Match to Dros ARALAR 1 Y18197

390	NPS0512	EST	3	360-501	AA438961	Human KIAA0160 gene
392	NPS0514	EST	3	1-237inv	AI064414	
393	NPS0515	EST	3	1-67inv	AA540712	809bp with AA440879 and AA440431.human cyclin G assoc. Kinase.
394	NPS0516	GNL	3	596-682inv	AF132145	Drosophila damage-specific DNA binding protein DDBa p127 subunit
395	NPS0517	EST	3	1-513	AF007159	Dros cDNA AF007159
396	NPS0518	EST	3	13-235inv	AI511691	
397	NPS0519	EST	3	1-61inv	AA264883	also AA392712. Alt splice. EP line AQ254720-lethal ?
398	NPS0520	EST	3	30-451	AA438399	821bp contig with AA439438. EP line AQ254604-lethal?
399	NPS0521	EST	3	600-627	AA440272	1324bp contig with AA438941. see also genomic AC013237
400	NPS0526	EST	3	65-483	AA264865	720bp EST contig with AA439894 and AI061812(end .Alt splice).
401	NPS0527	EST	3	1-246	AA263693	Dros Unknown AF132150
402	NPS0528	EST	3	37-160and, 379-47	AA698620	
403	NPS0529	EST	3	1-51inv	AA391350	
404	NPS0530	EST	3	100-499	AA392183	1194bp EST contig. Seq. sim. to human DNA J
405	NPS0531	EST	3	49-348 and 412-45	AA696390	Also AI387907. Seq. sim. to mammalian ubiquitin degradation fusion protein 1
406	NPS0532	EST	3	1-267inv	AA802961	816bp with AA817584 .Match to Dros. complete mRNA AF145690. EP line AQ073773-lethal?
407	NPS0533	EST	3	1-259inv	AA699045	1989bp EST contig.random slug cDNA25 protein
408	NPS534	EST	3	1-99bp	AA952055	1205bp with AA202358, AA202625 and AA951416. Siah binding protein 1(human)
409	NPS0535	EST	3	1-610inv	AA142266	
410	NPS0536	EST	3	52-534	AA696974	684bp EST contig with AA736106. See also genomic AC014104. Matches Human CGI-37 protein
411	NPS0537	EST	3	03-626 and 734-106	AI532170	1565bp with AI544333 and AI062662. Definite transcription factor, MTF-1
412	NPS0538	EST	3	442-569	AA567128	Match to mouse EST.EP line AQ025072
413	NPS0540	EST	3	1-79bp	AA950480	
414	NPS0541	GNL	3	1-360	AJ006772	Dros cyclin B3
415	NPS0542	EST	3	30-194 and 291-366	AA539625	882bp with AA202440 and AA390927. Dros Unknown AF132164
416	NPS0543	EST	3	1-255	AA951297	Also AI108147. Seq. sim. to mouse lysyl hydroxylase 3
417	NPS0544	EST	3	20-95 and 479-654	AA948996	1516bp EST contig. Seq. sim. to ZNF127-Xp [Homo sapiens]
419	NPS0546	EST	3	228-522	AA801928	
420	NPS0547	EST	3	14-98	AA951147	960bp EST contig with AA695598, AI109677 and AA540269. Poss. protein phosphatase.
421	NPS0548	EST	3	1-279	AA697191	1002bp with AA392404 and AA438791. Seq sim. to mitogen inducible gene mig-2 - human
422	NPS0549	EST	3	650-689	AI518422	
423	NPS1065	EST	3	27-689	AI535025	
424	NPS0551	EST	3	8-563	AA950826	
426	NPS0555	EST	3	411-582	AI109292	Poss heat shock protein
427	NPS0556	EST	3	1-279inv	AA202259	Same ESTs as NPS0483, different insertion point.
428	NPS0557	EST	3	131-647	AA142065	Also AI109234. poss. succinate semialdehyde dehydrogenase

430	NPS0559	GENO.	2	1-217	AC020266	At 18.6-19.2kb. Poss in space before gene coding for 669AA ORF.
431	NPS0560	GENO.	2	1-606	AC015232	At 2.5-3k. Intron of gene at 2-22k. 180AA ORF.
432	NPS0561	GENO.	2	1-169inv	AC013001	At 2.2-2.4k. 1.8kb before gene at 3-5kb. 183AAORF.
434	NPS0563	GENO.	2	1-849inv	AC014142	At 37.7-38.6k(odd gaps?). No good exons
435	NPS0564	GENO.	2	1-585	AC017578	6.3-6.8k. Space before EST contig coding for ORF with Seq sim to mam. NADH dehydrogenase
436	NPS0565	GENO.	2	1-505inv	AC015232	At 10.2-10.7k. Intron of gene at 2-22k. 180AA ORF.
437	NPS0566	GENO.	2	1-581	AC015178	At 8.6-9.2k. In intron gene at 4-18k. 244AA ORF. No database matches. Large ORFS on Rev too.
438	NPS0567	GENO.	2	1-637	AC017945	At 1-1.5k. ESTs inc AA949804 at 1.5-3.9k no matches
440	NPS0570	GENO.	2	1-662inv	AC015222	At 115.2-115.9K.Intron of gene represented by mRNA AF145688 on reverse strand.
441	NPS0571	GENO.	2	1-496inv	AC019859	At 9.4-9.9kb. In intron of gene coding for a 1121AA ORF. Seq. sim to membrane type metalloproteases.
442	NPS0572	GNL	2	1-559inv	AC007299	At 92145-92705. 348AA ORF at 104-91k.Intron of Dros GDI (see NPS0008) at 90658-93408
443	NPS0573	EST	2	50-397	AL062213	Only 1000bp long/EST AI532704 1158 EST contig matching 2303cDNA with 537AAORF
444	NPS0574	GENO.	2	1-479inv	AC020183	At 74.5-75kb. In intron of gene represented by ESTs inc.AA264274. Poss transcription factor.
445	NPS0575	GENO.	2	1-182inv	AC007257	At 132883-133065. EST AA735518 at 132600bp
446	NPS0576	GENO.	2	1-370	AC017289	At 6.1-6.5k. No good exons in this area. ESTs at 2.5k no matches.
447	NPS0577	GENO.	2	1-435	AC019995	At 12-12.4kb. In intron of gene at 22-4k. 946AA ORF. some sim. to other hypo. genes.
448	NPS0578	GNL	2	1-235	AC017330	At 7.3-7.6kb. 4bp overlap with Dros clathryn light chain AF055900
449	NPS0579	GENO.	2	1-328	AC013001	At 4.7-5k. Part of gene at 3-5k coding for 183AA ORF.
450	NPS0580	GENO.	2	1-110	AC017678	At 14.5kb. In intron of gene coding for 948AA ORF. Seq. sim to Mam. Microsomal triglyceride transport protein. ESTs confirm.
451	NPS0581	GENO.	2	1-472	AC019893	At 13.1-13.6. In intron of gene on rev. coding for 671AA ORF. Seq. sim to mouse sex-determination protein homolog Fem1a
452	NPS0582	GENO.	2	1-790	AC020017	At 18-18.8kb. In intron of gene at 25-5k. 369AA ORF.
453	NPS0584	GENO.	2	1-404	AC014987	At21.5-22k. 1kb 5' of gene represented by ESTs inc. AI257570
455	NPS0586	GENO.	2	1-518	AC017529	At 1.3-1.8kb. Poss. in space before gene coding for 360 AAORF at 3-19k (ESTs at 16-18k)
456	NPS0587	GENO.	2	1-324	AC015180	At 31.3-31.6kb. ORF on Rev. good match to transcription factors.
457	NPS0588	GENO.	2	1-325	AC015179	At 10-1.4kb. In intron of gene at 4-16kb(complement).1173AA predicted ORF. Good homology to human SEC24
459	NPS0590	GENO.	2	1-571inv	AC015233	At 9.3-9.9kb. No matches this area.
460	NPS0591	GENO.	2	1-455inv	AC014124	At 7.9-8.4kb.In intron of gene coding for 427AA protein. Similarity to alpha-amidating enzymes. ESTs at 7Kb
462	NPS0593	EST	2	1-295inv	AI532704	
464	NPS0595	GENO.	2	1-470	AC014326	At 1.5-2kb. Short sequence, no exons found.

465	NPS0596	GENO.	2	1-507	AC015424	At 115.4-115.9kb. In intron of gene at 101-129k, 1267AA ORF. EST matches at 119k
466	NPS0597	GENO.	2	1-260inv	AC018184	At 117.4-117.7kb.Space before 367AA ORF at 118k. Poss DNA binding protein.
468	NPS0599	GNL	2	1-615	U34383	At 2.6-3.2kb in AC017782. In intron of Dros Dopamine receptor DopR99B.
469	NPS0600	GENO.	2	1-27bp	AC014368	At 40kb. In intron of gene coding for 530AA ORF. Poss. DNA binding protein
470	NPS0601	GENO.	2	1-551	AC014496	At 1246-1804bp. In intron of gene represented by ESTs at 600bp-9kb.349AA ORF v.good sim. to coronin proteins.
471	NPS0603	GENO.	2	1-465inv	AC015064	At 60-60.5kb. Also AC006247. In intron of gene on rev. 574AA ORF. seq.sim to receptor tyrosine kinase.
472	NPS0604	GENO.	2	1-215	AC008370	At 56841-57049bp.1kb downstream of Dros mutL gene
473	NPS0605	GENO.	2	1-412inv	AC017529	At 8.7-9.1. In intron of gene at 3-19k, 360AA ORF. ESTs at 16-19k.
474	NPS0606	GENO.	2	1-559	AC017138	AT 28.7k-29.2k. Space before ESTs at 21-28k coding for 311AA ORF on reverse strand poss.retinaldehyde-binding protein
475	NPS0607	GENO.	2	1-474inv	AC020031	At 7.7-8.1kb. In intron of gene coding ofr 952AA ORF. Poss. WD repeat protein. ESTs confirm
477	NPS0609	GENO.	2	1-157	AC020227	At 116kb. Poss. iin space bewtween Dros. H3.3 gene and ESTs at 117k
479	NPS0611	GENO.	2	1-485inv	AC020007	At 31.1-31.5kb. In intron of gene coding for 1152AA ORF. Seq. sim to mam. sphingosine phosphate lyase 1. ESTs confirm.
480	NPS0612	GENO.	2	1-1145inv	AC014142	Same area as NPS0563. Diff. insert point. No good exons
482	NPS0614	GENO.	2	1-522inv	AC014423	At 55.1-55.6kb. Gene at 51k on rev , gene at 67k. nothing in between.
483	NPS0615	GENO.	2	1-325	AC020320	At 42.5-42.8kb. No good exons/ESTs found
484	NPS0616	GENO.	2	1-426inv	AC014073	At 5.6-6kb. in intron of gene coding for 800AA ORF. Good similarity histone acetyltransferase.(human)ESTs at 3k.
485	NPS0617	GENO.	2	1-527	AC019753	At 90-90.5k. In intron of gene at 81-104k. 720AA ORF. Seq sim to carboxypeptidases. EST confirms
486	NPS0618	GENO.	2	1-504inv	AC017381	At 60.6-61.1kb. Intron of gene coding for 1514AA ORF. Seq. sim to mammalianMAP/ERK kinase kinase 4. ESTs confirm
491	NPS0623	GENO.	2	1-333inv	AC015399	At 44.2-44.6. Space before gene represented by EST AI517547
492	NPS0624	GENO.	2	1-91bp	AC017666	At 5.7k. In intron of gene at 1-8k. 397AA ORF. Seq sim to mouse zinc finger protein 40. ESTs confirm
494	NPS0626	GENO.	2	1-548	AC015064	At 9.8-10.4kb. In intron of gene coding for 185AA ORF. No matches /ESTs
495	NPS0627	GENO.	2	1-120inv	AC017502	At 1166-1285bp. Space before gene represented by EST AA695901
496	NPS0628	GNL	2	1-408	AF132145	Same as NPS0516. diferent insert. point.
497	NPS0629	GENO.	2	1-559inv	AC007146	At 134370-135130bp. Full length cDNA at 135140-140020.ie space before this gene
498	NPS0630	GENO.	2	1-592	AC017643	At 40.8-41.4kb. In intron of gene at 32-45kb. 1058AA ORF. ESTs at 34 and 39k.
499	NPS0631	GNL	2	1-108inv	AF216532	AT 14kb in AC017347. In intron of Dros Migraine gene.
500	NPS0632	GENO.	2	1-284	AC020443	At 15.5-15.8.Intron of gene at 17-3k. Coding for 449AA ORF. Seq sim to Sad1 unc-84 domain protein 1 [Homo sapiens]. ESTs confirm
502	NPS0634	GENO.	2	1-522inv	AC019961	At 600bp. No good exons in this area. Dros tetraspanin 1 at 8-14k

503	NPS0635	GENO.	2	1-676bp	AC014142	At 24.8-25.5k. No good exons in this area. ESTs at 17k and 27k.
504	NPS0636	GENO.	2	1-541	AC017396	At 11.1-11.6kb. In intron of gene at 28-9k. coding for 378AA ORF. ESTs confirm.
505	NPS0637	GENO.	2	1-59inv	AC009197	At 90836-90894bp. 436AA orf at 86000-89400bp. EST AI513240 at 90600bp
506	NPS0638	UA	2			Matches various genomic sequences only one section not whole length?
507	NPS0639	GENO.	2	1-234	AC017928	At 85kb.In intron of gene at 95-70k coding for 1160AA ORF.
508	NPS0640	GENO.	2	1-31inv	AC014497	At 44kb. In intron of gene represented by ESTs at 42-45k. Diff. insert point to NPS0105, same ESTs.
509	NPS0641	GENO.	2	1-892inv	AC014076	Strange Gap in match. Poss. not correct genomic?
511	NPS0643	GENO.	2	1-197inv	AC020308	At 2-2.2k. In intron of gene at 2-16k.456AA ORF ESTs confirm.
512	NPS0644	GENO.	2	1-305	AC015089	At 36.1-36.4kb. Space before gene represented by EST AI543228 on the reverse strand.
513	NPS0645	GENO.	2	1-313	AC015089	At 35.7-36.1.15bp overlap with NPS0644
514	NPS0646	GENO.	2	1-530inv	AC017981	At 18.8-19.3kb In intron of gene at 1-21k coding for 915AA ORF. Some sim to mito carrier proteins.
515	NPS0647	GENO.	2	1-516	AC007452	At 98781-99300bp. in Intron of gene represented by EST AI517398. V.good sim. to ATPases
516	NPS0648	GNL	2	1-583inv	AC012981	At 14-14.5k. In intron of Dros DDP-1 DNA binding protein AJ238847. No flybase
517	NPS0649	GENO.	2	1-436	AC018631	At 52.6-53(n's) Incomplete sequence
518	NPS0650	GENO.	2	1-442	AC017382	At 1-1.4k. V close to previously published lethal insert which overlaps an EST.
519	NPS0651	GNL	3	1-536	AC014121	At 41.8-42.3kb. Space before Dros glut-1 on reverse strand,
520	NPS0652	GENO.	3	1-469inv	AC010120	At 45-45.4kb. In intron of gene represented by EST AI516046. Matches full length cDNA AF160941.
521	NPS1066	GENO.	3	1-417inv	AC013107	At 26.6-27kb. In intron of gene coding for 928AA ORF with seq sim to luciferases. ESTs confirm
523	NPS0655	GENO.	3	1-510	AC020444	At 7.7-8.2kb. gene on rev. v. sim to human splicing factor 2 at 9k onwards (ESTs confirm)
524	NPS0656	GENO.	3	1-527inv	AC020018	At 76.4-76.9kb. Space before gene coding for 573AA ORF with good sim. to potassium channel genes.
526	NPS0658	GENO.	3	1-408inv	AC012939	At 16.8-17.2kb. Predicted gene at 23- 20k (EST confirms) on rev. No matches in area of our insert.
529	NPS0661	GENO.	3	1-348inv	AC014810	At 45.4-45.7kb. 3bp overlap with EST AI534852. In gene coding for 1354AA ORF.
530	NPS0662	GENO.	3	1-463inv	AC018338	At 9.6-10.1kb. In intron of gene represented by ESTs inc. AA950503.
531	NPS0663	GENO.	3	1-150inv	AC013963	At 43.1-43.2kb. ESTs at 46k and 47k. No other matches.
532	NPS0664	GENO.	3	1-439	AC010662	At 109930-110370bp. 875AA orf on rev. Poss. hormone receptor. ESTs match this orf inc. AA735548
535	NPS0667	GENO.	3	1-461	AC018284	At 41.8-42.3kb. No good exons/ESTs here.
537	NPS0669	GENO.	3	1-544bp	AC013201	At 34.7-35.2kb. 400bp 5' to EST AI258331. No matches.
538	NPS0670	GENO.	3	1-530inv	AC019785	At 61.4-61.9kb.2kb 3' to the start of TOLLO receptor on rev.strand.
541	NPS0673	GENO.	3	1-513	AC015427	At 13-13.5kb. No exons/ESTs this area.
543	NPS0675	GENO.	3	1-611	AC015159	At 20.3-20.9kb. No good exons/ESTs in this area.

545	NPS0677	EST	3	455-586inv	AI944858	And AI946620 alt. splice. Genomic AC014457
547	NPS0679	GENO.	3	1-371inv	AC017381	At 57.3-57.7kb. Space before gene coding for 1514AA ORF. Seq, sim to mammalian MAP/ERK kinase kinase 4. ESTs confrim
548	NPS0680	GENO.	3	1-539inv	AC020108	At 7.8-8.3kb. No good exons /ESTs in this area.
549	NPS0682	GENO.	3	1-449	AC015358	At 9.4-9.9kb. 27bp overlap with EST AA264916(part of cluster at 9.5-8.3k). 287AA ORF with Good Hom to Rat tip49. TATA binding protein.
551	NPS0684	GENO.	3	1-480	AC017581	At 76.4-76.9kb. In intron of gene at 50-77k.2216AA ORF. Good sim to Mouse A-RAF kinase.
553	NPS0686	GENO.	3	1-515	AC014262	At 7.4-7.9kb. Probably in intron of gene represented by ESTs at 7 and 8K. No other matches.
554	NPS0687	GENO.	3	1-357	AC015138	At 50.7-51kb. In intron of gene coding for 1075AA ORF with some sim. to Human inhibitor of kappa light polypeptide gene enhancer in B-cells. EST at 48k.
555	NPS0688	EST	3	10-462, 518-end. in	AI541626	159AAorf, no hits. Genomic AC018224
557	NPS0690	GENO.	3	1-203	AC019753	At 50.8-51kb. No predicted genes in this area. No ESTs.
558	NPS0691	GENO.	3	1-202	AC019694	At 77.9-78.1kb. In gene on rev(ESTs at77 and 79) with good seq. sim to ribosomal protein L31.
559	NPS0692	GENO.	3	1-311	AC017810	At 21.3-21.6kb. Poss in intron of gene matching a yeast probable membrane protein. EST at 21.7k
560	NPS0693	GENO.	3	1-511	AC013175	At 6.1-6.6kb. Probably in first intron of TCP-1-eta chaperonin homologue. ESTs at 7K onwards.
562	NPS0695	GENO.	3	1-505	AC014850	At 1.8-2.3kb. No exons/ ESTs found.
563	NPS0697	GENO.	3	1-406	AC013939	At 35.4-35.8kb. In intron of gene represented by ESTs Inc AA817130.
564	NPS0698	GENO.	3	1-368inv	AC020332	At 32-32.3k In intron of gene coding for 445 AA ORF. Seq. sim. to NADH oxidoreductase genes.
565	NPS0699	GENO.	3	1-278	AC020248	At 11.2-11.5kb. In gene represented by ESTs at 11.2k AA979097.
566	NPS0700	GENO.	3	1-290inv	AC014875	At 49.5-49.8kb. Poss part of gene coding for a 377AA protein with some sim. to human CGI-88(ESTs at 53-58k also, may large gene).
568	NPS0702	GENO.	3	1-766inv	AC015159	At 5.4-6.2kb. Poss in intron of FTZ-F1 gene.
571	NPS0705	GENO.	3	168-497	AC009457	NPS0705 has no ECOR1 site yet doesn't match along whole length, first 168bp are unrelated to AC009457(or other matching clones)
573	NPS0707	GENO.	3	1-1306	AC019746	At 14.7-16k. Poss. on intron of gene coding for 991AA ORF. ESTs at 10k and 23k.
575	NPS0709	EST	3	47-125	AI531164	And AI518916.Genomic AC009344
576	NPS0710	GENO.	3	1-375	AC014674	At 20.2-20.5kb. In space before gene represented by EST at 19k.AI454966.
577	NPS0711	GENO.	3	1-322	AC013945	At 38.5-38.8kb. No exons/ESTs in this region
578	NPS0712	GENO.	3	1-67inv	AC014444	At 32.8-33.1kb. ESTs upto 30k. None in this area.
579	NPS0713	GNL	3	1-348inv	AF055900	Matches AC009383. Dros Clathryn matches AF055900, different insert point to NPS0578
580	NPS0714	GENO.	3	1-316inv	AC014674	At 15.2-15.5kb. EST AI297184 upto 15k
581	NPS0715	GENO.	3	1-511	AC010028	At 22575-2308bp 199AA orf. poss peptidoglycan recognition protein

WO 01/18547

PCT/GB00/03444

583	NPS0717	GENO.	3	1-490	AC014872	At 36.2-36.7kb. In intron of gene represented by ESTAI260669. ORF on rev, 350AA poss s-phase kinase.
584	NPS0718	GENO.	3	1-409	AC019531	At 61.5-61.9kb. In intron of gene coding for 1243AA ORF. ESTs matching whole ORF. Inc. AA941806
585	NPS0719	GENO.	3	1-705	AC015426	At 9.7-10.4kb. No exons/ESTs in this area
588	NPS0722	GENO.	3	1-480	AC013946	At 2.6-3.1kb. No good exons/ESTs
589	NPS0723	GENO.	3	1-114inv	AC014946	At 29.4-29.7kb. No good exons/ ESTs in this area.
590	NPS0724	GENO.	3	1-460inv	AC017270	At 88.2-88.6kb. Poss. in first intron ofgene coding for 1382AA orf on rev. ESTs confirm.
592	NPS0726	GENO.	3	1-300	AC019588	At 27.4-27.7kb. No exons/ESTs in this area.
593	NPS0727	GNL	3	1-184inv	AC009391	Also AC015336 and AC013186. Part of gene represented by ESTs inc AA201355. Match to Dros PTB. AF211191
594	NPS0728	GENO.	3	1-866inv	AC019542	At 31.9-32.8kb. 23bp 5' to EST AI404673.No other hits
595	NPS0729	EST	3	57-352inv	AC009537	At 24892-25196bp. orf matching ESTs AA735710 and AA942459. at 25220bp. V.like mannose 1-phosphate guanyltransferase
597	NPS0731	GENO.	3	1-443inv	AC019528	At 1.2-1.6kb. In intron of gene represented by EST AI515487. V.good match to 2-oxoglutarate deH.
598	NPS0732	GENO.	3	1-352inv	AC017582	At 15.5-15.9kb. In intron of gene coding for a 645AA ORF. No ests/matches.
599	NPS0733	GENO.	3	1-513inv	AC015353	At 12.9-13.4kb. ESTs at 11 and 18k. None in this area.
600	NPS0734	GENO.	3	1-600	AC014473	At 64.9-65.6k 900bp from start of gene represented by complete cDNA AF132183 on rev.
603	NPS0737	GENO.	3	1-371	AC020211	At 20.4-20.7kb ESTs at 19 and 21k. Probably in gene represented by these. 2050bp EST contig.471AA ORF.
604	NPS0738	GENO.	3	1-488	AC015395	At 44.5-45kb.Space before gene coding for 485AA ORF on reverse. Poss. membrane prot. ESTs at 40-41k.
606	NPS0740	GENO.	3	1-387	AC013175	At 18.6-19kb. Probably in space before gene represented by ESTs at 21k
607	NPS0741	GENO.	3	1-322inv	AC017856	At 14.4-14.8kb. no good exons here. ESTs at 11 and Dros ribosomal s12 gene at 19k.
608	NPS0742	GENO.	3	1-590inv	AC017582	At 34.9-35.5kb. gene upto 31k. EST at 37k. nothing at 31.5-37k.
609	NPS0743	GENO.	3	1-416	AC013197	At 13.2-13.6kb. Part of gene at 13-7k on rev. 558AA. No matches.
610	NPS0744	GENO.	3	1-504inv	AC019750	At 74.3-74.8kb. No good exons/ESTs in this area.
611	NPS0745	GNL	3	1-879inv	AC017132	At 35.5-36.4kb. 300 bp from start of Dros Proteasome gene X70304.
612	NPS0746	GENO.	3	1-443inv	AC018338	At 6.1-6.5kb. A few bps 5' to EST AA801928.
613	NPS0747	EST	3	1-231	AI517830	Also genomic AC014084
614	NPS0748	GENO.	3	1-473	AC014209	At 3.6-4kb. 20bp 5' to EST AI404170. No other matches.
615	NPS0749	GENO.	3	110-188	AC020211	At 17.4-17.6kb. just after a gene on plus strand. No ESTs until 19k
616	NPS0750	GENO.	3	266-439	AC019754	At 6-6.4kb. No good exons/ESTs in this area
618	NPS0752	GENO.	3	1-410	AC019753	At 122.4-122.8kb. No exons/ESTs this area.
619	NPS0753	GENO.	3	1-531inv	AC020031	At 57.3-57.8kb 200bp 5' to EST(AI402026) with seq. sim to rolyl 4-hydroxylase, alpha subunit.
620	NPS0754	EST	3	407-546inv	AI546104	Also genomic AC007577

621	NPS0755	GNL	3	1-462	AC014889	At 23.3-23.7kb. In intron of transforming acidic coiled-coil containing protein (tacc)AF146700.
622	NPS0756	GNL	3	1-145inv	AC017685	At 73.1-73.3kb. 300bp 5' to cDNA AF181626 coding for 1429AA protein, no database matches.
624	NPS0758	GNL	3	1-249	AC014071	From 1.7-1.9kb. In intron of Dros E78B, U01088.
625	NPS0759	GENO.	3	1-534inv	AC019869	At 9.7-102kb. ESTs at 5-8kand 17k. Nothing in this area.EST at 8k.has apoptosis gene similarities.
626	NPS0760	GENO.	3	1-557	AC019742	At 24.2-24.7kb. ESTs inc. AI297362 at 23.5-24.1kb.
628	NPS0762	GENO.	3	1-408	AC014368	At 6.8-7.2kb. 1.7kb 5' to EST AA392444
629	NPS0763	GENO.	3	1-566inv	AC017811	At 36.5-37.1kb. No good exons/ESTs this area.
630	NPS0764	GENO.	3	1-570	AC012807	At 14.7-15.3kb. No good exons/ESTs this area
631	NPS0765	GENO.	3	1-579	AC017241	At 40.8-41.3kb.in intron of gene for 500AA ORF. Good seq sim to Succinyl-COA:3-Ketoacid-Coenzyme A Tranferase precursor. Many ESTs confirm.
632	NPS0766	GNL	3	1-511	AC014110	At 2.5-3.1kb. In first intron of Dros Barbu gene AF132987.
633	NPS0767	GENO.	3	1-505	AC015229	At 55.8-56.3kb. Gene predicting 1429AA ORF at 40-70kb. Seq sim. to FAT tumour supressor.
635	NPS0769	GENO.	3	1-210inv	AC020260	At 61.8-62.1k. 15bp 3' to ESTs inc. AA952007. Part of gene coding for 591AA ORF on rev. Seq sim to P.walti rnp associated protein 55.
636	NPS0770	GENO.	3	1-317	AC009904	At 66179-66500. 111AA ORF on Rev, good match to Human DNA dir. RNA pol II NP_002687.1
637	NPS0771	EST	3	1-170	AI534756	Also genomic AC019950, at 126.4-126.6kb.EST doesn't match but may be in intron of tropomyosin II. lethal known 1989.
638	NPS0772	GENO.	3	1-317inv	AC019758	At 84.3-84.6kb. Space before ESTs on rev. at 82k. Good seq. sim to mouse nuclear body associated kinase 1b
639	NPS0774	GNL	3	109-606inv	AC019824	At 11.2-11.8kb. In intron of dros Bab-1 gene, AJ252082. No flybase.
640	NPS0775	GENO.	3	1-375inv	AC013046	At 19.3-19.7kb. In intron of gene coding for 859AA protein. Sim to other hypothetical genes. ESTs at 15-21k confirm.
641	NPS0776	GENO.	3	1-435inv	AC013969	At 3.3-3.8kb. ESTs at 7kb no hits in this area.
642	NPS0777	GENO.	3	1-790inv	AC018284	Overlap NPS0667 Diff. insert point.
643	NPS0778	GENO.	3	1-565	AC012986	Not a complete match.
644	NPS0779	GENO.	3	1-511inv	AC018173	At 13-13.5kb. In intron ofgene coding for 505AA ORF on rev. Good seq. sim to other hypothetical proteins. ESTs confirm.
645	NPS0780	GENO.	3	1-558	AC019746	At 2.9-3.4kb. No good exons/ESTs in this area.
646	NPS0781	EST	3	482-573bp	AI946243	No further hits. V. poor sequence
647	NPS0782	GENO.	3	1-507	AC014870	At 4.9-5.5kb.In intron of gene represented by ESTs inc. AI114091.
648	NPS0783	GENO.	3	1-30inv	AC019806	At 30.5kb. 2bp 5' to EST AA950675. Probable transcription factor.
649	NPS0784	GENO.	3	1-412	AC019585	At 54.9-55.3kb. Poss. in space before gene coding for 1796AA ORF with good seq.sim to endoplasmins/HSPs. ESTs confirm.
650	NPS0785	GENO.	3	1-492inv	AC015425	At 9.9-10.4kb. No good exons/ ESTs in this area.
651	NPS0786	GENO.	3	1-582	AC020454	At 3.3.5kb. No ESTs/exons this area.

652	NPS0787	GENO.	3	1-518	AC015075	At 4.8-5.3kb. No ESTs/ Exons in this area
653	NPS0788	GENO.	3	1-446	AC014674	At 34.6-35.1kb. 40bp 3' to ESTs inc.AA949585.
654	NPS0789	GNL	3	1-403	AF221066	At 18.6-19kb in AC015332. In intron ofDros Kurtz arrestin gene.
655	NPS0790	GENO.	3	1-525	AC020081	At 28.1-28.6kb.Prob. in intron of gene on rev.from 42-23k.good seq sim to branched chain ketoacid dehydrogenases. ESTs confirm
656	NPS0791	GENO.	3	1-589inv	AC013164	At 19.7-20.3kb. ESTAA141118 at 19.3 K may be in intron of gene represented by this EST (end of genomic clone).
657	NPS0792	GENO.	3	1-528inv	AC019758	At 83.3-83.8kb. In space beffore gene represented by EST at 82.2k AI533899.
659	NPS0794	GENO.	3	1-756	AC014423	At 33.2-34kb. In intron of gene coding for 256AA orf with good similarity to myosin phosphatase regylatory subunit.
660	NPS0795	GENO.	3	1-630	AC019651	At 13.9-14.5kb. 1kb 3' to ESTs inc. AA141648.
661	NPS0796	GENO.	3	1-162	AC015227	At 10.5 -10.6kb. In intron of gene represented by EST AA820684.No good matches.
662	NPS0797	GENO.	3	1-509inv	AC013024	At 1.1-1.5kb. No ESTs/ exons in this area
663	NPS0798	GNL	3	1-182inv	AC018013	At 15.6-15.8kb. Space before Dros receptor like guanylate cyclase, U23485. Diff. insert point to NPS0075.
664	NPS0799	GENO.	3	1-541	AC019758	At 11.4-11.9kb. No good exons/ESTs in this area.
665	NPS0800	GENO.	3	1-633inv	AC012837	At 8.9-9.6kb. In space before gene at 7k coding for 250 AA ORF. ESTs confirm.
666	NPS0801	GENO.	3	1-460inv	AC010052	At 8.3-8.7kb. 14bp before cDNA AF160911. Sim to dead box proteins.
667	NPS0802	GENO.	3	1-443	AC017671	At 26.3-26.8kb. ESTs at 25k Large ORF at 27.4k?
668	NPS0803	GENO.	3	1-524inv	AC019498	At 86-600bp. Plus 5k from AC019567. 4bp 5' to ESTAA142181 part of gene coding for 350AA ORF.
669	NPS0804	GNL	3	1-537inv	AC018013	At 17-17.5kb. Space before HLH gene U38238. Diff. insert point to NPS0064.
670	NPS0805	GENO.	3	1-459inv	AC013958	At 27.6-28kb. Either in 3'region or Dros NMDA receptor. or in space before Na/K cotransporter.
671	NPS0806	GENO.	3	1-371	AC014084	At 48.8-49.1kb. Dros proteasome gene at 44k, ESTs at 52k, nothing in this area.
672	NPS0807	GENO.	3	1-551inv	AC017890	At 42.1-42.6k. cDNA ending at 43.7k. ORFs/ESts at 37k. No hits this area.
674	NPS0809	GENO.	3	1-515inv	AC014087	At 2.5-3kb. 300bp 5' to EST AI062268(no db matches.)
675	NPS0810	GENO.	3	1-513	AC018127	At 10.4-10.9kb,. In intron of gene represented by ESTs inc AI456639.2770bp contig no db matches.
676	NPS0811	GENO.	3	1-549inv	AC017731	At 11.7-12.3kb EST at 12.9k incAI108782, no DB matches
677	NPS0812	GENO.	3	1-339inv	AC014784	At 70.3-70.6kb. In intron of gene represented by ESTs at 71-68k. Poss ATPsynthase. EP line too, lethal?
678	NPS0813	GENO.	3	1-582inv	AC014994	At 5.1-5.6kb. ESTs at 7-12k. no other exons/matches.
679	NPS0814	GENO.	3	1-323	AC017623	At 16.68-17kb. Poss. in intron of gene represented by ESTAI063960 at 16641bp.
680	NPS0815	GENO.	3			small genomic fragment
681	NPS0816	GENO.	3	1-722inv	AC006402	At 43.6-44.4kp. 34bp 5' to ESTAI109098. No db matches
682	NPS0817	GENO.	3	1-860inv	AC015221	At 5.9-6.7kb. No exons/ESTs this area.
683	NPS0818	GENO.	3	1-570inv	AC017270	At 82.4-83kb. In intron of gene coding for 1382 AA ORF on rev. ESTs confirm

684	NPS0819	GENO.	3		AC020026	At 40.2-40.7kb. In intron of gene coding for 612AA orf, ESTs at 41-44k.Poss transcription factor.
685	NPS0820	GNL	3	1-22inv	AC018284	At 32k. 200bp before LK6 knase U76378.
686	NPS0821	GENO.	3	1-378inv	AC015351	At 12-12.3kb. No exons / ESTs in this area
687	NPS0822	GENO.	3	1-541inv	AC020316	At 6.4-6.9kb. ESTs at 9.3k, none in this area.
688	NPS0823	GENO.	3	1-427inv	AC012807	At 7.7-8.1kb. No good exons/ESTs this area
689	NPS0824	GENO.	3	1-157inv	AC020316	At 7.1-7.2kb. ESTs at 9.3k, none in this area.
690	NPS0825	GENO.	3	154-408inv	AC020539	At 45.6-46kb. EST at 44k, no other matches.
691	NPS0826	GENO.	3	1-441inv	AC017691	At 48-48.4kb. No exons/ESTs in this area.
692	NPS0827	GENO.	3	1-686	AC014819	At 19.1-19.8kb. 100bp 3' to EST AI294326, part of gene coding for 389AA ORF on rev.
693	NPS0828	GENO.	3	1-927inv	AC014840	AT 54.7-55.6kb. 1 kb 3' to start of dros spazle gene on rev. strand.
695	NPS0830	GENO.	3	1-201	AC017740	At 99.2-99.4kb. 1kb from start of dros Dofl in rev. strand. AJ010642.
696	NPS0831	GENO.	3	1-114inv	AC020304	At 48-48.1kb. Poss. in intron of gene at 47-73k.898AA ORF.ESTs from 65k, prob. phospholipase.
698	NPS0833	GENO.	3	1-786	AC015358	At 7.3-8k.In gene on rev. coding for 310AA ORF.Seq. Sim. to HumanTATA binding protein interacting protein 49 kDa. EST cluster confirm this.
699	NPS0834	GENO.	3	1-574inv	AC012776	At 9.9-10.5kb. No exons/ESTs this area.
700	NPS0835	GNL	3	1-621inv	AC020154	At 34.8-35.4kb. In intron of Dros FAX gene, U21685. EP lines known.
702	NPS0837	GENO.	3	1-469	AC017267	At 21.1-21.6kb. 1kb 5' to anothre p-element claiming to be in Hearty, no match on AC014595.
703	NPS0838	GENO.	3	1-963	AC012791	At 12.1-13kb. In intron of gene represented by EST AA439616. No db matches.
705	NPS0840	GENO.	3	1-754	AC020432	At 3.4-4.2. EST at 2.3kb. no other hits.
707	NPS0842	GENO.	3	1-989inv	AC019820	At 27.5-28.4kb. In intron of gene represented by ESTs from 20-37k. V. good sim. to rat casein kinase.
708	NPS0843	GENO.	3	1-183inv	AC020006	At 96.1-96.3kb. In intron of gene represented by EST at 95-97k.AI546104. poor match to WD-40 repeat prot.
709	NPS0844	GENO.	3	1-304inv	AC019806	At 396-709bp. 1bp overlap with gene at 708-1538. Good seq. sim. to phosducin like genes.ESTs confirm.
710	NPS0845	GENO.	3	1-855	AC015138	At 63.3-64.1kb. Large ORF in this area, no DBmatches. No ESTs.
711	NPS0846	GENO.	3	1-825inv	AC020332	At 29.2-30kb. In gene coding for 445 AA ORF. Seq. sim. to NADH oxidoreductase genes.
712	NPS0847	GENO.	3	1-798	AC015342	At 42.1-42.9kb. No good exons/ESTs in this area.
713	NPS0848	GNL	3	1-797inv	AC013201	At 7.4-8.2kb. 500bp 5' to Dros SIN mRNA.AF151390.
714	NPS0849	GENO.	3	1-491inv	AC017905	At 52.9-53.4kb. No exons/ESTs in this area.
715	NPS0850	GENO.	3	1-1013	AC017721	At 71.8-72.8kb. In area of predicted gene with sim. to cytosolic phospholipase A2 beta. EST at 74k.
716	NPS0851	GENO.	3	1-902	AC014376	At 122.6-123.5kb. In intron of gene coding for 320AA ORF . V. good seq. sim to human phosphoribosyl pyrophosphate synthetase. ESTs confirm.

717	NPS0852	GENO.	3	1-64bp	AC020327	At 26.8kb. 60bp 5' to EST AI542698. Other ESTs from 27-41k(introns). may be part of bumetanide-sensitive Na-K-2Cl cotransporter
720	NPS0855	GENO.	3	1-110inv	AC019562	At 120.3-120.4kb. In space before gene coding for 286AA ORF on rev. Seq. sim. to Protein Phosphatase2C
721	NPS0856	GENO.	3	1-1070inv	AC018185	At 15-16kb.Either space before Dros MPCP gene on rev or before gene represented by EST at 18k.
722	NPS0857	GENO.	3	1-765	AC013945	At 36.3-37.1kb. No good exons/ESTs in this area
723	NPS0858	GENO.	3	1-568	AC017811	At 41.4-41.9kb. No good exons/ESTs this area.
724	NPS0859	GENO.	3	1-580	AC014351	At 17.4-18kb. In intron of gene at 21-9kb. 432AA ORF, seq. sim. to mouse nucleotide binding protein 2.
725	NPS0860	GNL	3	1-403inv	AC012937	At 38.9-39.3kb. Probably in intron of Dros polypyrimidine tract binding protein (PTB).AF211191.
726	NPS0861	GNL	3	1-416	AC017132	At 23.2-23.7kb. In intron of Dros fatty acid desaturase.U73160. diff. insert point to NPS0048.
727	NPS0862	GENO.	3	1-52bp	AC010070	At 900bp. Part of gene represented by EST AI515487. V.good match to 2-oxoglutarate deH.
728	NPS0863	GENO.	3	1-490	AC020332	At 33.6-34.1kb. No exons/ESTs this area.
729	NPS0864	GENO.	3	1-1153inv	AC019826	At 12.4-13.5kb. In space after EST at 10k or before Dros Brahma at 14.8k
730	NPS0865	GNL	3	1-1144	AC017132	At 22.6-23.7. Intron of FAD U73160.Diff insert point to NPS 0861/NPS0048.
731	NPS0866	GENO.	3	1-858	AC015441	At 26.6-27.4kb. In 1st intron of gene coding for342AA ORF on rev. Seq. Sim. to yeast autophageosis prot. ESTs confirm.
732	NPS0867	GENO.	3	1-832inv	AC018100	At 36.3-37.2kb. In intron of gene coding for 354AA ORF. EST at 39k confirms.
733	NPS0868	GENO.	3	1-532	AC012947	At 99.4-99.9kb.No good EST/Exons in this area.
735	NPS0870	GENO.	3	1-1145	AC017676	At 45-46kb. AC017592 matches end of AC017676. 1kb 3' to EST at 44k.
736	NPS0871	GNL	3	1-447inv	AC015301	At 4.2-4.7kb. Probably in first intron/promoter of Dros SF1 gene on rev.AJ243904.
738	NPS0873	GENO.	3	1-885	AC019753	At 141.3-142.2kb. Just 3' to ESTs at 140-141k inc.AI387574. Seq. sim. to human UDP-glucose 4-epimerases.
739	NPS0875	GENO.	3	1-1083	AC012873	At 30.7-31.8kb. No exons/ESTs this area.
740	NPS0876	GENO.	3	1-1796	AC014884	At 9.9-11.7kb. No exons/ESTs this area.
741	NPS0877	GENO.	3	1-819	AC015220	At 1.5-2.3kb. No exons/ESTs this area.
742	NPS0878	GENO.	3	1-1003	AC014889	At 16.8-17.9kb. Probably in promoter of gene represented by cDNA AF132153. Probable endopeptidase.
743	NPS0879	GENO.	3	1-384	AC019683	At 385-768bp. EST /Gene sequence at 2kb. not perfect Match. Cytoplasmic dynein/annexin 9.
745	NPS0881	GENO.	3	1-519	AC012727	At 52-52.5kb. In intron of Dros anon 1A3 at 53.3kb. AF005843. ESTAA439438 confirms.
747	NPS0883	GENO.	3	1-99bp	AC019746	At 14.7kb. No exons/ESTs this area.
748	NPS0884	GENO.	3	1-423	AC020248	At 9.9-10.4kb. In intron of gene represented by ESTs at 1.6kb. inc AI517276. No db matches.
749	NPS0887	GENO.	3	1-495	AC010059	And start of AC015427. At 1.1-1.6kb. No exons/ESTs this area.
753	NPS0889	GENO.	3	1-556inv	AC017585	At 81.2-81.7kb. Poss. in intron of gene represented by ESTs at 80 and 82k. Some seq. sim. to nucleoporins.

754	NPS0890	EST	3	1-925	AC018045	At 26.7-27.6kb(end). 17bp overlap with EST AC018045.1435bp EST contig, no good db matches.
755	NPS0891	GENO.	3	1-1125	AC017250	At 26-27kb. 300bp 5' to EST AI944859.
756	NPS0892	GENO.	3	1-1137	AC014884	same area as NPS0876, diff. insert point
757	NPS0893	GENO.	3	1-534	AC013945	At 965-1581bp. No good exons/ESTs this area.
758	NPS0894	GENO.	3	1-527inv	AC019746	At 7.9-8.4kb. No exons/ESTs in this area.
759	NPS0896	GENO.	3	1-646	AC010005	At 37123-37769bp. ORF related to Human polypeptide chain release factor ERF1
760	NPS0897	GENO.	3	1-93bp	AC015358	At 5.8-5.9kb. Intron of gene represented by EST AI257536.
761	NPS0898	GENO.	3	1-1044	AC010005	At 80366-81430bp. ORF related to FRG1 from Fugu(muscular dystrophy related)
762	NPS0899	GENO.	3	1-1345	AC015076	At 25.6-26.9kb. 23bp overlap with EST AA141054, same as NPS0371, diff. insert point.
763	NPS0900	GENO.	3	1-597inv	AC014084	At 124.8-125.4kb. In intron of Dros gene represented by cDNA AF145694. 447AA ORF. Prob. adenylosuccinate synthetase 1
764	NPS0901	GENO.	3	1-577	AC019995	At 63.9-64.5kb. In intron of gene represented by EST AI402854. no db matches.
765	NPS0902	GENO.	3	1-940	AC020333	At 2-3kb. In intron of gene on rev, 402AA. Good seq. sim. to Ubiquitin carrier proteins. EST confirm.
766	NPS0903	GNL	3	1-1131	AC017132	At 24.3-24.4kb. 100bp 3' to start of Dros FAD .U73160. Diff. insert point to NPS0048.
767	NPS0904	GENO.	3	1-687inv	AC020260	At 60-4-61.1kb. In intron of gene represented by ESTs at 57-61.8kb. Part of gene coding for 591AA ORF on rev. Seq sim to P.walti rnp associated protein 55.
768	NPS0905	GENO.	3	1-510	AC017784	At 69.7-70.2kb. ESTs at 65 and 75k. No hits in this area.
770	NPS0907	GENO.	3	1-113inv	AC014365	At 4.1-4.2kb. 200bp 5' to ESTs inc. AI542840. V. good seq. sim to myeloid/lymphoid or mixed lineage-leukemia translocation to 10 homolog. MAF10.
771	NPS0908	GENO.	3	1-1116inv	AC017721	At 73.3-74.4kb.In area of predicted gene with sim. to cytosolic phospholipase A2 beta. EST at 74k.
773	NPS0910	GENO.	3	1-727inv	AC014325	At 22.9-23.6kb. May be in intron of gene on rev. coding for 1586 AA ORF with seq. sim. to ATP binding proteins.
774	NPS0911	GENO.	3	1-1010	AC015160	At 28.6-29.6kb. No good exons/EST in this area.
775	NPS0912	GENO.	3	1-1426	AC019772	At 0.6-2kb. No exons/ESTs in this area.
776	NPS0913	GNL	3	1-403	AC019850	At 86-86.4kb. In intron of dros protein disulphide isomerase gene.U18973.
777	NPS0914	GENO.	3	1-1111inv	AC014588	At 19.3-20.5kb. 35bp 5' to ESTAI543592.
778	NPS0915	GENO.	3	1-499	AC017495	At 23.9-24.4kb. ESTs at 21k, none in this area.
779	NPS0916	GENO.	3	1-371inv	AC014994	At 600-900bp.ESTs at 7-12k. no other exons/matches.
781	NPS0918	EST	3		AI514396	also genomic AC013006. At 4.5-5.5kb.
782	NPS0919	GNL	3	1-118	AC018013	At 15.5-15.6kb. Space before Dros Guanylate cyclase at 14.2kb on rev. U23485. Diff. insert point to NPS0075.
783	NPS0920	GENO.	3	1-176	AC017329	At 33.8-33.9kb. In intron of gene coding for 1500AA ORF from 25k with good sim. to UNC51 kinase. EST confirm(see NPS0383)
784	NPS0921	GENO.	2	1-537INV	AC006073	At 32.8-33.3kb. In intron of gene coding for 246AA. No db hits, no ESTs.

785	NPS0922	GENO.	2	1-720	AC004299	At 83.9-84.7kb.In intron of gene represented by EST AI533754
786	NPS0924	GENO.	2	1-599	AC004115	At 13.6-14.2kb. In gene represented by ESTs AI520524 and AI945841. No db matches.
787	NPS0925	GENO.	2	1-581	AC004716	At 72.4-73kb.In intron of gene coding for 355AA protein. 51777-83843bp. No database matches
788	NPS0926	GENO.	2	1-628	AC005889	At 36-36.6kb. No exons/ ESTs in this area.
790	NPS0928	GNL	2	1-86inv	AF165153	In intron of Dros DGP-1, developmental gene. See genomic AC004296
791	NPS0929	GENO.	2	1-573inv	AC004306	With AC004766inv. In intron of gene represented by ESTs AI238344, AI108292 and AI533605. At 52k-67k
792	NPS1077	GENO.	2	1-648	AC006472	At 71.6-72.2kb.In intron of gene represented by EST AI542461, part of 1506AA ORF. Poss. transcription factor.
793	NPS0931	GENO.	2	1-463inv	AC007175	At 67-67.4kb. In Intron of gene represented by ESTs inc AA941489.1647bp EST contig. No good db matches.
794	NPS0932	GENO.	2	1-519	AC006073	At 40.7-41.2kb. No exons/ ESTs in this area.
796	NPS0935	GNL	2	1-307inv	AC004423	In intron of Dros AKAP see also NPS288
797	NPS0936	GNL	2	1-412inv	AF167578	Space before Dros Septin 5. See genomic AC005448
798	NPS0937	GENO.	2	1-478	AC004313	At 39-39.4kb. Poss. in 3'UTR of gene coding for 355AA protein at 39.7-40.7.Weak seq. simi. to potassium channel gene
799	NPS0938	GENO.	2	1-489	AC004641	At 125.4-125.8kb. No exons/ESTs in this area
800	NPS1078	GENO.	2	1-558inv	AC004306	At 57.4-57.9kb. In intron of gene of gene represented by ESTs AI533605 and AI238344.
802	NPS0941	GENO.	2	1-544	AC005334	At 73.3,73.8kb. ESTs at 74650bp (AA803646, AI518976, AI108114) Sequence similarity to U5 snRNP
803	NPS0942	GNL	2	1-201	AC004154	Part of Dros NTPase different insert point from NPS28
804	NPS0943	GNL	2	1-524	AC004766	At 60.5-61kb. 130bp 3' to start of Dros Igloo gene on rev.S72579.Gap 43 like protein.
805	NPS0944	GENO.	2	1-621	AC004361	At 21.4-22kb. 338AA ORF at 19-32k, no db hits.
806	NPS0945	GENO.	2	1-569	AC020283	At 16.6-17.1kb. ESTs at 15k and 21k none in this region.
808	NPS0947	GENO.	2	1-233inv	AC005750	At 71.8-72kb. 630AA orf at 60-77k no hits. No ESTs either.
809	NPS0948	GENO.	2	1-525	AC005269	At 43.8-44.3kb. 274AA ORF at 40-50k,complement.Seq. sim to human TGFB inducible early growth response 2
810	NPS0949	GENO.	2	1-531	AC005554	At 19.8-20.3kb. Part of gene represented by EST AI402921. Part of large ORF matching mammalian fatty acid synthase.
812	NPS0952	GENO.	2	1-498inv	AC005894	At 63-63.4kb. No good exons/ESTs in this area
813	NPS0954	GNL	2	1-320	AC004564	At 34.6-34.9kb. ESTs at 30k match Dros syndecan,U03282.In large intron going into AC019924.
814	NPS0956	GENO.	2	1-429	AC005716	At 72.2-72.6kb.EST at 69-71k, probably part of this gene.
815	NPS0958	GENO.	2	1-71bpinv	AC011662	At 211.6kb.100 bp 5' to ESTAI108521 at 211-226k.
816	NPS1079	GENO.	2	1-75bp	AC004758	In intron of gene coding for a 945AA protein at 87648-113518. Strong sequence similarity to Human retinoblastoma binding protein 2

818	NPS0963	GENO.	2	1-512inv	AC004758	At 90-90.5k. In intron of gene coding for 1079AA ORF with good seq. sim. Human retinoblastoma binding protein 2. EST at 94k confirms.
819	NPS0964	GENO.	2	1-54bp	AC004334	At 19.8kb.In space b/n ESTs at 18k and those at 21k.
820	NPS966	GENO.	2	1-557inv	AC005149	In intron of gene coding for 424AA protein at 70149-97938. No database matches.
821	NPS0968	GENO.	2	1-202	AC005333	At 94.2-94.4kb. Also AC018171. No good exons/ESTs this area.
822	NPS970	GENO.	2	1-534inv	AC005334	In intron of gene coding for 309AA protein at 64276-77888bp. No database matches
823	NPS0971	GENO.	2	1-438	AC006421	At 7-7.4kb. Also AC005449. No good exons/ESTs this area.
824	NPS0972	GENO.	2	1-524INV	AC005443	At 23.3-23.8kb. No good exons/ESTs this area.
825	NPS0973	GENO.	2	1-492	AC005650	At 5.8-6.2kb. Poss. in intron of gene coding for 2245AA ORF represented by ESTAI947230 at 6.6k.Seq. sim to human uridine phosphorylase.
826	NPS0974	GENO.	2	1-535	AC005889	At 18.8-19.3kb. No good exons/ESTs this area.
827	NPS0975	GENO.	2	1-47bpinv	AC005130	At 13.6kb. In intron of gene represented by ESTs at 7-28k inc AA949050(2072bp contig).
829	NPS0977	GNL	2	1-100, 146-499	AF097364	Drosophila Drongo gene
831	NPS0979	GENO.	2	1-256inv	AC004722	At 32.6-32.9kb.Space before gene coding for 834AA protein at 33470-40630. Sequence similarity to bromodomain containing proteins.ESTs confirm.
832	NPS0980	GENO.	2	1-406	AC003054	At 25.2-25.6kb. Poss. in intron of gene at 14-46k, 538AA ORF. EST at 18-44k.
834	NPS0982	GENO.	2	1-460	AC004280	At 48.5-48.9kb. ESTs at 47.8k and 49.5k. No db matches.
835	NPS985	GENO.	2	1-178inv	AC001661	In intron of gene coding for 300AA protein at 30647-46841. Weak sequence similarity to Mouse surfeit gene
836	NPS0986	EST	2	1-602	AI533769	See also genomic AC005269. Other ESTs no DBmatches.
837	NPS987	GENO.	2	1-562	AC004362	No good predicted exons in this area.
838	NPS0988	GENO.	2	1-521	AC004370	At 54.6-55.1kb. 3bp 3' to EST AI124332 on rev. No db hits.
840	NPS0991	GENO.	2	1-535inv	AC005447	At 34.1-34.7kb. In intron of gene coding for a 802AA protein at 28-49kb.EST at 33, 40 and 48k.
841	NPS0992	GENO.	2	1-342	AC019901	119.6-119.9kb. No good exons/ESTs in this area
842	NPS0993	GENO.	2	1-512	AC005454	At 65.6-66.1kb. In space before gene coding for a 399AA protein at 66-68kbp. ESTs confirm. Seq. sim. to mitochondrial carrier protein genes.
843	NPS0994	GENO.	2	1-515inv	AC005130	At 42.5-43kb. No good exons/ESTs in this area
845	NPS0997	GENO.	2	1-565inv	AC012753	At 62.4-62.7. Space before gene represented by ESTs inc. AI294250. Seq. sim. to ribosomal protein L30.
846	NPS0998	GENO.	2	1-568	AC005889	AT 66.7-67.3kb. 3bp overlap with gene coding for1365AA protein at 67-74kb. ESTs AI106939 and AI296430 confirm.
847	NPS0999	GENO.	2	1-503	AC004370	AT 52.9-53.4kb.In intron of gene represented by EST AI124332.
848	NPS1000	GENO.	2	1-620	AC004351	At 36.3-36.9. In gene coding for 676AA protein at 34511-37955bp. Sequence similarity to mouse LUN gene.
849	NPS1001	GENO.	2	1-519inv	AC004766	At 11.2-11.7kb.In intron of gene coding for 1467AA protein at 0.8-18kbp. 3 bp overlap with EST at 11.7.Seq.sim. to Drosophila Lipase 3.

850	NPS1002	GENO.	2	1-80inv	AC006247	In gene coding for 805AA protein at 160506-163420bp.Sequence similarity to Mammalian Valyl tRNA synthetase.
851	NPS1003	GENO.	2	1-370inv	AC006574	At 13.7-14kb. In intron of gene coding for a 1131AA protein at 20-1k. No database matches. ESTs at 16-18k.
853	NPS1005	GENO.	2	1-535inv	AC005447	At 0.3-0.8kb. Also AC019827. 1kb 3' to Dros Anillin and ESTs.
854	NPS1006	GENO.	2	1-581inv	AC005643	At 30.5-31kb.In intron of gene coding for 593AA protein at 39-20kbp. Poss. calcium channel. EST at 31.5-33k.
856	NPS1009	GENO.	2	1-77inv	AC004532	At 13.8kb. 20bp Overlap with EST AI260111. 889bp EST contig. Seq. sim. to multivitamin transporter.
857	NPS1010	GENO.	2	1-496inv	AC019704	At 11.5-12kb. Between ESTs at 10k and ESTs at 13k.
858	NPS1012	GENO.	2	1-483	AC004423	At 4.5-5kb.In intron of gene coding for 604AA protein at 3-8.8k.Sequence similarity to C.elegans AL021481 gene.ESTs confirm gene
860	NPS1013	GNL	2	1-560	AC005811	Intron of Dros Steroid receptor L06423 FTZ-F1B
861	NPS1016	GENO.	2	1-596	AC005653	At 46.8-47.4kb.EST upto 43k. full length cDNA at 51k, nothing in between.
862	NPS1017	GENO.	2	1-539	AC004516	At 60-60.5k. Part of gene at 63-47k coding for 732AA ORF. Poss. exonuclease. EST at 54k.
863	NPS1019	GENO.	2	1-505inv	AC005285	At 119.7-120.2kb. In intron of gene coding for1142AA protein at 116605-128877bp. Sequence similarity to Guanine nucleotide exchange genes.
864	NPS1021	GENO.	2	1-504inv	AC020029	At 96.6-97.1kb. In intron of gene coding for 395AA ORF at 91-118k. Seq. sim . to inwardly rectifying calcium channels.
865	NPS1022	GENO.	2	1-191	AC005643	At 11.5-11.7kb. In intron of gene represented by ESTs inc. AI238833.
866	NPS1023	GENO.	2	1-468	AC004642	At 22.8-23.2kb. 100bp 3' to ESTAI519819. Part of gene on rev.1159AA ORF.Poss. lysophosphatidic acid acyltransferase.
867	NPS1024	GENO.	2	1-578inv	AC005749	At 11.3-11.8kb. Part of gene coding for 1481AA protein at 159-11694bp. Good Seq. sim. to JNK-binding protein JNKBP1Mouse. ESTs confirm.
868	NPS1025	GENO.	2	1-598	AC020254	At 2.6-3.2kb. ESTs at 1and 8k.
869	NPS1027	EST	2	1-40bp	AI257015	661bp EST contig. See also genomic AC004340.
871	NPS1029	GENO.	2	1-198	AC004375	At 30.5-30.7kb. ESTs at 32-34k. No exons/ESTs at 30k area.
872	NPS1030	GENO.	2	1-316	AC005472	At 78.5-78.8kb.Intron of gene represented by ESTs at 78-81k inc. AI258704. No db hits.
873	NPS1031	GNL	2	1-495inv	AC004154	At 11.6-12.1kb. 21bp overlap with Dros. geranylgeranyl transferase type II,AF133269.
874	NPS1032	GENO.	2	1-116	AC004328	At 22.6-22.7kb.In intron of gene coding for 1218AA protein at 2-28k. Has been predicted from Dros genomic AL035311 and has similarity to mouse BOP1. Many ESTs confirm.
875	NPS1033	GENO.	2	1-581inv	AC005112	At 33.7-34.2k. In intron of gene coding for 407AA protein at 3680-49217(complement). No database matches
876	NPS1034	GENO.	2	1-506inv	AC004367	At 2.8-3.3kb Space before gene coding for 387AA protein at 3496-35348bp. No database matches. ESTs at 7k
877	NPS1036	GENO.	2	1-411	AC005472	At 78.9-79.3kb.Intron of gene represented by ESTs at 78-81k inc. AI258704(different point of insertion to NPS1030)

879	NPS1038	GENO.	3	1-291	L49405	At 35.3-35.6kb. No good exons/ESTs this area.
880	NPS1039	EST	3	1-250	AI543199	1127bp contig with AI544292.See also genomic AC004658.
881	NPS1062	GNL	3	1-376	AC007757	Also matches Dros. EST AA951801.Part of Dros Elongin B (diff. insert point to NPS79)
882	NPS1044	GENO.	3	1-597	AC006091	At 102.6-103.2kb. In intron of gene represented by EST AA246373 at 100-104k. No db matches.
883	NPS1045	GENO.	3	1-498	AC005720	At 163.5-164kb. In intron of gene represented by ESTs AI260814 and AI947168.
885	NPS1049	GENO.	3	1-486	AC004713	At 16.4-16.8kb.In intron of gene coding for 575AA protein at 10435-29297. EST at 12-15k AI108650. Weak similarity to CDC36 genes.
886	NPS1050	GENO.	3	1-544	AC005813	No good predicted exons in this area.
887	NPS1051	GENO.	3	1-549	AC006936	At 95.5-96kb. No good exons/ESTs in this area
888	NPS1052	GNL	3	1-306inv	AJ236864	Dros. putative adenosine kinase. Homozygous deletion has no phenotype.
889	NPS1053	GENO.	3	1-579inv	AC005720	At 79.4-79.9kb.In intron of gene represented by ESTs at 78-83k.AI544292 and AI543199.
890	NPS1056	GENO.	3	1-191inv	AC004266	AT 68.4-68.6kb. In intron of gene coding for 931AA at 70-50kb. Seq. sim. to C. elegans Zinc finger protein. ESTs AI259457 and AI519745 at 61-67k.
891	NPS1059	GENO.	3	1-264	AC006936	At 132.3-132.5kb. No ESTs/Exons this area.
892	NPS1063	GNL	3	488-536	AI062190	EST comes from Dros ferrochelatase
893	NPS1064	GENO.	3	1-1068	AC019758	At 30-31k.Part of gene at 35-29k. 900 AA ORF.Seq sim to Human transcriptional co-activator CRSP150, AF104256. ESTs confirm.
894	NPS1076	EST	2	101-597	AI388606	1249bp contig with AI258281 and AI258326
895	NPS1081	EST	2	1-491	AI533993	AC005714
896	NPS1082	GENO.	2	1-475inv	AC005714	At 99630-10010bp. 1065AA ORF at 95-118k. EST AI295594. Similarity to mammalian RAD50/ poss. channel
897	NPS1083	GENO.	2	1-461	AC004375	At 31.1-31.5kb. different insert point NPS1029.Closer to EST at 32k.
898	NPS1084	GENO.	2	1-507inv	AC018075	At 16.3-16.8kb.Between ORF at 15k (seq. sim. to pre-mRNA splicing SR protein rA4) or Dros ERCC2 at 17.4k. EST at 15.8k.
899	NPS1085	GENO.	2	74-502	AC007413	At 22(n's)-22.4kb. Area of transposon insertions.
900	NPS1086	GENO.	2	225-525	AC015186	At 33.2-33.5kb. No good exons/EST in this area.
901	NPS1087	GENO.	2	1-521	AC017673	At 3.3-3.8kb. In gene on rev coding for335AA ORF. Mouse polycystic kidney disease-related protein.
902	NPS1088	GENO.	2	1-378	AC017676	At 35.2-35.5kb. In intron of gene represented by ESTs at 32-40k, inc AI542779. 2023bp EST contig . 666AA ORF. No db matches.

CLAIMS

1. A screening assay for identifying compounds which have a physiological effect on an organism, the assay comprising the steps of:

a) reacting a test compound with a protein encoded by an essential gene comprising a sequence selected from the group consisting of SEQ ID Nos. 1-902, specific fragment thereof, or homologue thereof, from the organism; and

b) detecting any modulatory effect the compound has on the protein.

2. The screening assay according to claim 1 wherein the sequence is selected from the group consisting of SEQ ID Nos. 430-783 and 899-902.

3. The screening assay according to any preceding claim wherein the effect on the protein is a negative modulation.

4. The screening assay according to any preceding claim wherein the assay is a ligand binding assay for detecting the effect the compound has on the ligand binding of the protein

5. The screening assay according to any one of claims 1 to 4 wherein the assay is a functional activity assay for detecting the effect the compound has on the functional activity of the protein.

6. The screening assay according to claim 5 wherein the functional activity assay is selected from the group consisting of kinase assays; protein phosphatase assays; adenyl cyclase assays; guanylyl cyclase assays; phosphodiesterase assays; nucleosidease assays; protease assays; protein secretion and/or import assays; nuclease assays; DNA metabolism assays; transcription factor assays; apoptosis assays; calcium utilisation assays; receptor/ion

channel assays; and G protein assays.

7. A compound having modulatory activity on a protein encoded by an essential gene, as identified by an assay according to any preceding claim.

8. A pesticidal formulation comprising a compound according to claim 7, together with a pesticidally acceptable excipient.

9. Use of a compound having modulatory activity on a protein encoded by an essential gene as identified by an assay according to any one of claims 1 to 6 or derivative or analogue thereof as a pesticide.

10. A pesticidally active compound identified by an assay according to any one of claims 1 to 6 and further tested for its ability to kill pests.

11. Use of a pesticidally active compound according to claim 10, in conjunction with other pesticides, herbicides and agriculturally usual auxiliaries as crop protection material.

12. A method of selectively modulating activity, in an organism, of a protein encoded by an essential gene comprising a sequence selected from the group consisting of SEQ ID Nos. 1-902 or a specific fragment thereof, or homologue thereof, comprising administering a compound that selectively modulates activity of the protein in the organism.

13. The method according to claim 12, wherein the selective modulation in activity of the protein has the result of substantially eliminating or severely reducing the activity of the protein, as compared to the activity of the protein without modulation.

14. The method according to claims 12 or 13, wherein the compound modulates the activity of the protein and has a minimal modulatory effect on other proteins of the organism.

15. The method according to any of claims 12 to 14, wherein the modulation in activity of the protein has the effect of being lethal or semi-lethal to the organism.

16. A method of modulating activity, in an organism, of a protein encoded by an essential gene comprising a sequence selected from the group consisting of SEQ ID Nos. 1-902 or a specific fragment thereof, or homologue thereof, comprising administering a compound, that selectively modulates activity of the protein, to an organism, and wherein the ability of the protein to modulate the activity of the protein is determined by:
- exposing the protein which has been produced by a genetically engineered cell expressing the protein, with the compound for a period of time;
- measuring the activity of the exposed protein using a ligand binding or functional activity assay; and
- comparing the activity of the exposed protein with an activity of a control protein which has not been exposed to the compound, so that compounds that modulate the protein activity are identified.

17. The method according to claim 16, for selectively modulating activity, in an organism, of a protein, further comprising the steps of:
- exposing a further cellular protein(s) of the organism to the compound for a period of time;
- measuring the activity of said further protein(s) using an assay(s) appropriate for such a purpose; and
- comparing the activity of said exposed further cellular protein(s) with an activity of a control protein(s) which has not been exposed to the compound, so that compounds

that substantially do not, or minimally modulate said further cellular protein(s) activity, are identified.

18. A method of identifying compounds having a potentially pesticidal activity caused by modulation of a protein encoded by an essential gene comprising a sequence selected from the group consisting of SEQ ID Nos. 1-902 or a specific fragment thereof, or homologue thereof, which comprises;

- obtaining the protein by heterologous expression of the essential gene in a host cell;
- employing the protein in an assay according to any one of claims 1 to 6 for detecting a compound which displays modulatory activity on the protein; and
- testing the compound which displays modulatory activity on the protein for its pesticidal activity on an organism.

19. A compound identified by the method according to claim 17 as having pesticidal activity.

20. Use of a compound according to claim 19 as a pesticide.

21. A pesticidal formulation comprising a compound according to claim 19 identified as having pesticidal activity, together with a pesticidally acceptable excipient

22. A method for the production of a pesticidal composition comprising identifying a compound that displays pesticidal activity using the method according to claim 18 and mixing the compound identified, or a derivative, or an analogue thereof, with a pesticidally acceptable carrier.

23. An isolated polynucleotide fragment comprising a sequence selected from the group consisting of SEQ ID Nos.430-783 and 899-902, a fragment thereof, or a homologue thereof.

24. An essential gene comprising a sequence selected from the group consisting of SEQ ID Nos.430-783 and 899-902, a fragment thereof, or a homologue thereof.

25. An isolated polynucleotide which hybridises under stringent conditions to a polynucleotide fragment selected from the group consisting of SEQ ID Nos. 430-783 and 899-902 or a fragment thereof.

26. Use of an isolated polynucleotide fragment comprising a sequence selected from the group consisting of SEQ ID Nos. 430-783 and 899-902, a fragment thereof, or a homologue thereof to identify and facilitate isolation of an essential gene.

27. Use of a polynucleotide fragment selected from the group consisting of SEQ ID Nos.430-783 and 899-902 or a fragment thereof, to identify and facilitate isolation of homologous sequences from other organisms.

28. Use of a polynucleotide fragment selected from the group consisting of SEQ ID Nos.430-783 and 899-902 or a fragment thereof, to identify and facilitate isolation of genes, from other organisms comprising homologous sequences.

29. An essential gene comprising a sequence selected from the group consisting of SEQ ID Nos.430-783 and 899-902, a fragment thereof, or a homologue thereof.

30. An expression vector comprising the essential gene according to claim 29.

31. An expression vector according to claim 30 comprising one or more control sequences capable of directing the replication and/or expression of an operatively linked essential gene.

32. A prokaryotic or eukaryotic host cell comprising the expression vector according to either of claims 30 or 31.

33. A method of producing a polypeptide comprising culturing a host cell according to claim 32 under conditions permitting expression of the polypeptide.

34. A polypeptide produced by the method of claim 33.

35. Use of the polypeptide according to claim 34 in an assay for detecting compounds which modulate activity of the protein.

36. Use of a polypeptide expressed by an essential gene comprising a sequence selected from the group consisting of SEQ ID Nos. 1-902, or a fragment, or a homologue thereof, in a pesticide screening assay for identifying a compound which modulates activity of the polypeptide.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 March 2001 (15.03.2001)

PCT

(10) International Publication Number
WO 01/18547 A1

(51) International Patent Classification[7]: G01N 33/68, C07K 14/435

(21) International Application Number: PCT/GB00/03444

(22) International Filing Date:
6 September 2000 (06.09.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
9921009.8 7 September 1999 (07.09.1999) GB

(71) Applicant *(for all designated States except US)*: **THE UNIVERSITY COURT OF THE UNIVERSITY OF GLASGOW** [GB/GB]; Gilbert Scott Building, University Avenue, Glasgow G12 8QQ (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **DAVIES, Roger, Wayne** [GB/GB]; University of Glasgow, IBLS Division of Molecular Genetics, Anderson College, 56 Dumbarton Road, Glasgow G11 6NU (GB). **KAISER, Kim** [GB/GB]; University of Glasgow, IBLS Division of Molecular Genetics, Anderson College, 56 Dumbarton Road, Glasgow G11 6NU (GB). **YANG, Ming, Yao** [CN/GB]; University of Cambridge, Department of Genetics, Downing Street, Cambridge (GB).

(74) Agents: **McCALLUM, William, Potter** et al.; Cruikshank & Fairweather, 19 Royal Exchange Square, Glasgow G1 3AE (GB).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *With international search report.*

— *Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/18547 A1

(54) Title: ESSENTIAL GENES AND ASSAYS RELATING THERETO

(57) Abstract: The present invention relates to screening assays for compounds which have a physiological (e.g. harmful or beneficial) effect on proteins identified as being essential. The inventors have identified nucleic acid sequence from genes encoding proteins which are thought to be essential, the lack of expression of which leads to a lethal or semi-lethal phenotype and have described assays which may be conducted to screen for suitable compounds. The present invention also relates to novel sequence per se.

DECLARATION AND POWER OF ATTORNEY FOR PATENT APPLICATION

Attorney Docket No. 9013-44

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled ESSENTIAL GENES AND ASSAYS RELATING THERETO,

the specification of which

[ ] is attached hereto

OR

[ ] was filed on 6 September 2000 as United States Application No. or PCT International Application Number PCT/GB00/03444 and was amended on ________ (if applicable).

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in Title 37 Code of Federal Regulations, § 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, § 119(a)-(d) or § 365(b) of any foreign application(s) for patent or inventor's certificate, or § 365(a) of any PCT International application which designated at least one country other than the United States of America, listed below and have also identified below any foreign application for patent or inventor's certificate, or of any PCT International application having a filing date before that of the application on which priority is claimed.

9921009.8	United Kingdom	09/07/99	[X] Yes [ ] No
Number	Country	MM/DD/YYYY Filed	Priority Claimed
			[ ] Yes [ ] No
Number	Country	MM/DD/YYYY Filed	Priority Claimed
			[ ] Yes [ ] No
Number	Country	MM/DD/YYYY Filed	Priority Claimed

I hereby claim the benefit under Title 35, United States Code, § 119(e) of any United States provisional application(s) listed below.

Application Number(s)	Filing Date (MM/DD/YYYY)
Application Number(s)	Filing Date (MM/DD/YYYY)

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s) or § 365(c) of any PCT international application designating the United States of America, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application(s) in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application (37 C.F.R. § 1.63(d)).

Appln. Serial No.	Filing Date	Status Patented/Pending/Abandoned
Appln. Serial No.	Filing Date	Status Patented/Pending/Abandoned
Appln. Serial No.	Filing Date	Status Patented/Pending/Abandoned

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

POWER OF ATTORNEY: As a named inventor, I hereby appoint the practitioners associated with the Customer Number provided below to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith, and direct that all correspondence be addressed to that Customer Number:

Customer Number 20792

Send correspondence to: Kenneth D. Sibley
Myers Bigel Sibley & Sajovec
Post Office Box 37428
Raleigh, NC 27627

Direct telephone calls to: Kenneth D. Sibley
(919)854-1400

Facsimile: (919) 854-1401

1-00
Full name of (first/sole) inventor: ROGER WAYNE DAVIES

Inventor's Signature: [signature] Date: 7/3/02

Residence: University of Glasgow, IBLS Division of Molecular Genetics, Anderson College
56 Dumbarton Road, Glasgow G11 6NU, UNITED KINGDOM

Citizenship: British GBN

Post Office Address: University of Glasgow, IBLS Division of Molecular Genetics, Anderson College
56 Dumbarton Road, Glasgow G11 6NU, UNITED KINGDOM

2-00
Full name of second inventor: KIM KAISER

Inventor's Signature: [signature] Date: 23.04.02

Residence: University of Glasgow, IBLS Division of Molecular Genetics, Anderson College
56 Dumbarton Road, Glasgow G11 6NU, UNITED KINGDOM

Citizenship: British GBN

Post Office Address: University of Glasgow, IBLS Division of Molecular Genetics, Anderson College
56 Dumbarton Road, Glasgow G11 6NU, UNITED KINGDOM

Full name of third inventor: 3-00 MING YAO YANG

Inventor's Signature: M Yang Date: 2 MAY 2002

Residence: University of Cambridge, Department of Genetics Downing Street, Cambridge, UNITED KINGDOM

Citizenship: Chinese / British CNX

Post Office Address: University of Cambridge, Department of Genetics Downing Street, Cambridge, UNITED KINGDOM

Full name of fourth inventor: _____

Inventor's Signature: _____ Date: _____

Residence: _____

Citizenship: _____

Post Office Address: _____